The Goddard
Distributed Active Archive Center

NOAA/NASA Pathfinder AVHRR Land Data Set

User's Manual

Version 3.1

September 1994

ACKNOWLEDGMENTS

The requested form of acknowledgment for any publication in which the Pathfinder AVHRR Land Data are used is:

"Data used by the authors in this study include data produced through funding from the Earth Observing System Pathfinder Program of NASA's Mission to Planet Earth in cooperation with National Oceanic and Atmospheric Administration. The data were provided by the Earth Observing System Data and Information System, Distributed Active Archive Center at Goddard Space Flight Center which archives, manages, and distributes this data set."

This User's Manual should be cited as:

Agbu, P.A., and M.E. James. 1994. The NOAA/NASA Pathfinder AVHRR Land Data Set User's Manual. Goddard Distributed Active Archive Center, NASA, Goddard Space Flight Center, Greenbelt.

The editors of this document wish to thank John G.R. Townshend, Chairman of the AVHRR Pathfinder Land Science Working Group, for his help during the preparation of this manual.

THE PATHFINDER AVHRR LAND DATA SET USER'S MANUAL

This document provides a comprehensive guide to the NOAA/NASA Pathfinder AVHRR Land data (hereafter, all references to "Pathfinder" relate specifically to the Pathfinder AVHRR Land project and data sets).

Section 1 presents an overview of the various data products that constitute the Pathfinder data sets. If you already have Pathfinder data and are ready to begin reading and analyzing the data, you may go directly to Section 4, "Reading the Data Set," but be sure to review Section 2, "Data Set Details," and Appendix C, "Product Notes," if you are not familiar with the data details.

These data are available from the EOSDIS Distributed Active Archive Center at Goddard Space Flight Center (Goddard DAAC). If you want information on how to access the data archive to order data, order documentation, or to request assistance, refer directly to Section 3, "Archive and Data Access."

Section 5 presents a variety of information in the form of "Frequently Asked Questions."

CONTENTS

1. Data Set Overview

The Pathfinder AVHRR Land data sets are global, land surface data derived from the Advanced Very High Resolution Radiometers (AVHRR) on the NOAA/TIROS operational meteorological satellites (NOAA–7, –9, and –11) that provide a continuous daily and composite data set from July 1981 through the present. These data, when complete, will enable studies of global vegetation and surface characteristics over more than a decade. This section provides a high level overview of the Pathfinder data that are available from the Goddard DAAC.

SECTION 1 CONTENTS

1.1 Pathfinder Data Sets

There are four types of Pathfinder AVHRR Land data sets.

- Daily Data Set — daily, global, 8 km, terrestrial data
- Composite Data Set — 10-day, global, 8 km composites of daily data
- Climate Data Set — 1-degree NDVI product for climate modeling
- Browse Images — images for Daily and Composite Data Set selection

The Pathfinder project is also generating a data set for AVHRR calibration research that contains calibration data records extracted from the AVHRR orbital data. An ancillary data file with elevation and geolocation information is also provided for use with the Daily and Composite Data Sets.

Table 1-1 presents a summary of the size, structure, and format of the Pathfinder data products. Many of the Pathfinder data sets contain multiple geophysical parameters in each file. These parameters are organized as "stacked" two dimensional arrays, which are referred to in this document as **data layers**.

Also mentioned in Table 1-1 is the product format. All Pathfinder data, except the Calibration Extraction product, are stored in the Hierarchical Data Format (HDF) (Brown et al., 1993), which is the standard data format for the early Earth Observing System Data and Information System (EOSDIS). HDF is discussed in detail in Section 4, "Reading the Data," and Appendix E, "NCSA Hierarchical Data Format (HDF) Basics."

Table 1-1
Product Summary

Data Product	Data Layers Per File	Dimensions (columns x rows)	Format	Approximate Size Compressed/Uncompresse
Daily Data Set	12	5004 x 2168	HDF (SDS)	35 MB / 228 MB
Composite Data Set	12	5004 x 2168	HDF (SDS)	35 MB / 228 MB
Ancillary Data File	4	5004 x 2168	HDF (SDS)	18 MB / 50 MB
Climate Data Set	1	360 x 180	HDF (SDS)	12 KB / 66 KB
Browse Images	2	625 x 271	HDF (RIS8)	95 KB / 342 KB
Calibration Extracts	n/a	n/a	NOAA 1B	5 MB

1.1.1 Daily Data Set

The Daily Data Set contains global, 8 km, terrestrial data mapped to an equal area projection. Geophysical parameters contained in the data set include Normalized Difference Vegetation Index (NDVI) (Goward et al., 1991), cloud and quality control flags, solar and scan geometry, reflectances derived from AVHRR channels 1 and 2, brightness temperatures derived from AVHRR channels 3, 4, and 5, and date and hour of observation. Data over oceans, large inland water bodies, and in areas of twilight have been masked out.

These data are derived directly from the AVHRR level 1B orbital data using the procedures described in Appendix B. There is one file per day for the entire Pathfinder processing period (June 25, 1981, to present).

The Daily Data Set is useful for studies of many terrestrial variables (e.g., vegetation, temperature, snow cover) and for producing a variety of composite data sets.

1.1.2 Composite Data Set

Because the Daily Data Set contains clouds, a Composite Data Set has been created to enable land surface studies. The Composite Data Set is similar to the Daily Data in structure and is derived from the Daily Data Set; however, the process of compositing removes much of the cloud cover present in the Daily Data Set (Holben, 1986). The Composite Data Set contains the same geophysical parameters as the Daily Data Set, and these data are mapped to the same global, 8 km, equal area projection as the daily data. There are three composites per month for each year of data. The first composite of each month is for days 1 to 10, the second composite is for days 11 to 20, and the third composite is for the remaining days.

To generate the Composite Data Set, 8 to 11 consecutive days of data are combined, taking the observation for each 8 km bin from the date with the fewest clouds and atmospheric contaminants as identified by the highest NDVI value. Only data from the Daily Data Set that are within 42 degrees of nadir are considered in generating the composite.

The Composite Data Set is particularly useful for studies of temporal and interannual behavior of surface vegetation and for developing surface background characteristics for use in climate modeling. However, if a completely cloud free background is required, further compositing may be necessary.

1.1.3 Climate Data Set

This data set contains global, mean, cloud free, NDVI data at 1-degree resolution for each 8- to 11-day composite period. It is derived from the Composite Data Set, and there are 36 climate data files for each year.

The Climate Data Set is intended primarily for use in Global Climate Models (GCM), simple biosphere models (Sellers et al., 1986), and other global time series studies.

1.1.4 Browse Images

Browse Images are generated from the Daily and Composite Data Set to aid in scene selection. The Browse Images are produced by subsampling every eighth pixel of every eighth line of the full resolution daily and composite data files, and reducing 16-bit data to 8-bit. A daily Browse Image consists of two data layers, Channel 2 reflectance and Channel 4 brightness temperature, and is useful in determining cloud extent and areas of missing data. Each 10-day composite Browse Image consists of an NDVI and a Channel 4 brightness temperature layer and is useful for identifying areas of residual cloud contamination.

Because of their rescaling and subsampling, these images are for data selection only and are not recommended for data analysis.

1.1.5 Ancillary Data File

The Pathfinder Ancillary File contains land and sea flags, elevation, and bin center latitude and longitude for global 8 km bins that have all been coregistered to the same 8 km equal area projection as the daily and composite data.

There is a single, invariant ancillary data file for the entire Pathfinder time period and it is primarily used to assist in interpreting the daily and composite data. The land/sea mask and elevation in this file are the same used in the generation of the Daily Data Set; however, this file should not be confused with the ancillary data sets that are part of the processing inputs.

1.1.6 Extracted Calibration Data

To support AVHRR Calibration research, selected information is extracted from the satellite orbital data. This includes calibration coefficients, gains and offsets, blackbody/deep space view information, and instrument (baseplate) temperature.

Calibration data are extracted from approximately 6 Global Area Coverage (GAC) orbits per month. The orbit that passes over the equator at 0 degree longitude every 5–7 days is chosen. This results in data over different surface types (ocean and desert) within each orbit and different solar and scan geometry among orbits. These data are in the NOAA 1B 10-bit packed format as described in the "NOAA Polar Orbiter Data User’s Guide" (Kidwell, 1991), and are roughly 5 MB each.

1.2 Data Set Origins

The input to the Pathfinder data processing is more than 70,000 orbits of AVHRR GAC 1B data from the NOAA polar orbiting satellites with the afternoon ascending node equator crossing times. With current definitions of satellite data processing levels, these data would be referred to as level 1A; however, the GAC orbital data from the NOAA Polar Orbiters are referred to as 1B for historical reasons.

The AVHRR instrument onboard the NOAA series satellites (TIROS–N/NOAA 6–12) provides daily coverage of the Earth in 4 or 5 spectral bands at a nominal resolution of 1 km. Because the 1 km resolution data are too voluminous to be captured daily, the data are subsampled and averaged onboard and then transmitted to central receiving stations as GAC data with a nominal resolution of 4 km providing full global coverage.

Pathfinder input data commence with NOAA–7, which was launched in June 1981. This was the first AVHRR instrument with five channels, and it is the five channel AVHRR instruments that are used in producing the Pathfinder data sets. The additional channel provides better cloud discrimination and is useful for determining Sea Surface Temperature, which will be produced in a separate Pathfinder effort.

The AVHRR instrument's 110.8 degrees cross-track scan equates to a swath of about 2700 km. The orbital period is about 102 minutes, and there are 14 orbits per day with a repeat cycle of approximately 14 days.

A more detailed, comprehensive description of the NOAA series satellites, the AVHRR instrument, and the AVHRR GAC 1B data can be found in the "NOAA Polar Orbiter Data User's Guide" (Kidwell, 1991), which can be obtained from NOAA's National Environmental Satellite Data and Information Service (NESDIS) at

NOAA/NESDIS
Satellite Data Services Division
Room 100
Princeton Executive Square
5267 Allentown Road
Camp Springs, MD 20746

Phone: (301) 763-8402
Fax: (301) 763-8443

Internet: sdsdreq@pes.sdsd.ncdc.noaa.gov

The input ancillary data are described in Section 5, "Frequently Asked Questions."

1.3 Processing Features

An AVHRR Land Science Working Group (LSWG) consisting of university and federal agency scientists was chartered to recommend the content of and generation procedures for the Pathfinder data. At the request of the Land, Atmospheres, and Oceans Working Groups, an AVHRR Calibration Working Group (CWG) was formed to determine the best available calibration methodology to use. The members of the Land Science and Calibration working groups are listed in Appendix A. The procedures and algorithms used to produce these data are based on the recommendations of these groups.

The LSWG identified three areas where the Pathfinder processing should incorporate improved scientific methods. These processing features, which are unique to the Pathfinder data processing, include

- **Calibration** — Calibration of the visible channels uses new coefficients that incorporate many vicarious calibration techniques and aircraft underflight calibration (Rao, 1993a). The thermal channels are calibrated using methods that take into account the nonlinear response of the detectors (Rao, 1993b).
- **Navigation** — Errors of up to 20 km are observed in Earth locations calculated from predicted satellite positions, and there are known errors in the NOAA-provided Earth locations in the GAC data. For these reasons, a navigation system using the University of Colorado's orbital model (Baldwin and Emery, 1993) and the Goddard Space Flight Center/SeaWiFS geolocation system (Patt and Gregg, 1994) were incorporated into the Pathfinder processing.
- **Atmospheric Correction** — The visible and near-infrared channels of the AVHRR are corrected for Rayleigh scattering and ozone absorption. No corrections for aerosol or water vapor effects have been applied.

Appendix B provides a detailed description of the algorithm and processing methods used to produce the Pathfinder data set.

Several small variations in the Pathfinder processing system have occurred during the processing of the Pathfinder data sets due to errors being discovered, which necessitated some software modifications. These changes were determined by the Pathfinder Processing Team and the LSWG to have minimal effect on the data product, and were implemented beginning with the 1988 data processing. When all data (1981–1996) are processed, reprocessing of the 1987 data will be carried out to ensure data consistency. Appendix D, "Processing Changes," provides a chronological summary of data set processing inconsistencies. When a change is made in processing, the DAAC will issue online Processing Change Notes, update the user manual, and communicate any changes to the scientific community.

1.4 Validation

Before beginning production of the Pathfinder data, the Pathfinder processing algorithm was validated by processing test periods and checking against correlative data. Checks of all data layers were made comparing the Pathfinder values with related and precursor AVHRR data sets such as data from the First ISLSCP Field Experiment CD-ROM (Strebel et al., 1992) and the NOAA Global Vegetation Index (Goward et al., 1993; Kidwell, 1990).

A Validation Scientist, working with the scientific community, is responsible for overseeing product processing and quality, recommending modifications to quality control procedures, and working with the science community when questions of Pathfinder data quality arise. This scientist performs visual examination and statistical analysis of randomly selected images (daily and composite) and runs histogram analyses and multichannel cross evaluations on Pathfinder data. Temporal profiles of three surface types (e.g., forest, agricultural land, and desert) are maintained for seasonal comparison with all Pathfinder data produced.

1.4.1 Operational Quality Control

There are three forms of ongoing, operational quality control of the Pathfinder data; automated checks, visual inspection, and product and metadata checks.

Automated Quality Checks — A few validation checks have been built into the

Pathfinder data processing (see Section 2, "Quality Control Flags" in "Data Set Details"). Automated quality checks are made for consistency in fields such as date and satellite or scan times. Geophysical values are checked to see that they are within a reasonable range. Certain anomalies may exist in the data set because of conditions inherent in the input data — for example, missing scan lines or orbits, incorrect or incomplete calibration coefficients — and many of these data are flagged in the QC data layer.

Visual Data Inspection — The Pathfinder Operations Team maintains product quality control by performing a visual check on every layer of the daily and composite files. During this "gross error check" the team adds quality comments indicating areas of missing data, bad scans, and small navigation errors. These quality comments are added to the file and can be viewed in the DAAC's online Information Management System (IMS) (see Section 3, "Archive and Data Access") or by extracting the comments from the HDF file. An example of QC comments is shown in Table 1-2. Any serious errors will result in investigation by the Pathfinder Processing Team and Validation Scientist and subsequent reprocessing.

Table 1-2
Sample QC Comment Annotations

July 22, 1986, data	qc_comment=	Missing an orbit over western South America and another partial orbit over southeastern Canada. QC flags appear over India, the Middle East, and western Russia but do not seem to affect the channel data.
	qc_date=	May 06, 1994
July 23, 1986, data	qc_comment=	Missing a small area of data over Yucatan Peninsula to Cuba. Missing a scan in northwest Africa, southwest Asia. A bad scan is in central Brazil.
	qc_date=	June 07, 1994

The Daily and Composite Browse Images do not go through visual quality checks. The Climate Data Set is quality controlled, with quality control comments entered only when an exceptional condition is present in the Composite Data Set file from which the climate file was generated.

Product and Metadata Checks — Data checking and validation of the Pathfinder data files are carried out before data are transferred to the archive. Data archival involves extracting the metadata (descriptive text about the data sets), making sure that all of the expected metadata are present in the file, and copying the data files into the DAAC mass storage system. If the verification process fails, the DAAC staff notifies the Pathfinder Operations Team to recheck and correct their files before resending. The Transmission Control Protocol/Internet Protocol (TCP/IP) monitors the data file transfer through the network and performs a bit-by-bit integrity check on the data file. Before data are archived or distributed, the DAAC staff performs quality checks to ensure that files copied from disk to tape remain complete and correct.

2. Data Set Details

The Pathfinder data sets are useful for a variety of applications. It is important to understand the details of the Pathfinder data before using them. This section provides details needed to use the data effectively for scientific applications. For more detail on the processing procedures, see Appendix B, "Processing Algorithm Description." For information on reading and scaling the data, refer to Section 4, "Reading the Data Sets."

Additional important information on the data set details can be found in Appendix C, "Product Notes."

SECTION 2 CONTENTS

2.1 Daily Data Set

2.1.1 Data Layer Details

The Daily Data Set contains 12 layers of information (Table 2-1). The Daily Data Set files, which are in the HDF Scientific Data Set (SDS) format, also contain several types of descriptive, textual information referred to as metadata. The file metadata are described in Section 4.1.2, "Metadata and Tags." Product notes, which are important caveats and features of the data set, are provided in Appendix C. Users are encouraged to review this information before beginning any research using these data.

All of the data layers shown in Table 2-1 are stored as scaled 8- or 16-bit unsigned integers. Information on scaling the data to geophysical units is presented in Section 4, "Reading the Data Sets."

Table 2-1
Daily and Composite Product Data Layers

Layer	Units	Range
NDVI	—	–1 to +1
CLAVR flag	—	0 – 31
QC flag	—	0 – 31
Scan Angle	radians	–1.047197 to +1.047197 (~±60°)
Solar Zenith Angle	radians	0 – 1.396256 (0 – 90°)
Relative Azimuth Angle	radians	0 – 6.2832 (0 – 360°)
Ch1 Reflectance	%	0 – 100
Ch2 Reflectance	%	0 – 100
Ch3 Brightness Temperatures	Kelvin	160 – 340
Ch4 Brightness Temperatures	Kelvin	160 – 340
Ch5 Brightness Temperatures	Kelvin	160 – 340
Day of Year	DDD.HH	1– 366.23

Note that the AVHRR GAC 1B data used to create the Pathfinder data have a 10-bit resolution. However, Pathfinder data processing performs all calculations as 32-bit floating point integers, then convert them to 8- or 16-bit integer values in the output products. In all cases, the trailing 2 or 3 decimal places in data values do not necessarily represent increased data accuracy. The Pathfinder processing retains output data as 16-bit rather than 10-bit to unburden users from additional processing. When using the data, bear in mind that statistical results from calculations of geophysical values showed no significant trends regarding the trailing decimal places.

- **NDVI —** The Normalized Difference Vegetation Index (NDVI), which is related to the proportion of photosynthetically absorbed radiation, is calculated from atmospherically corrected surface reflectances (R) from the visible and near infrared AVHRR channels as

  $$\frac{R_{CH2} - R_{CH1}}{R_{CH2} + R_{CH1}}$$

  where R_{CH1} is the land surface reflectance in the visible wavelengths (0.58–0.68 micrometer) and R_{CH2} is the land surface reflectance in the reflective infrared wavelengths (0.725–1.1 micrometer). The principle behind this is that Channel 1 is in a part of the spectrum where chlorophyll causes considerable absorption of incoming radiation, and Channel 2 is in a spectral region where spongy mesophyll leaf structure leads to considerable reflectance (Tucker, 1979; Jackson et al., 1983; Tucker et al., 1991).

- **CLAVR Flags** — The Pathfinder data includes cloud flags produced by the NOAA Clouds from AVHRR (CLAVR) algorithm (Stowe et al., 1991). The CLAVR algorithm performs a series of threshold and uniformity tests on a 2 x 2 array of pixels. If 1–3 pixels in the array are flagged as cloudy, all four pixels are flagged as mixed. Otherwise the array is flagged as clear or cloudy. Zero values mean no decision was made because of missing data. The specific thresholds and tests used by CLAVR are shown in Table 2-2, and the clear, mixed, and cloudy flag values are shown in Table 2-3.

Table 2-2
CLAVR Cloud Detection Test Conditions, Stowe et al. (1991).

	Cloud Test	Thresholds and Values for 2 x 2 Array
C3AR	Channel 3 Albedo Restoral	The Ch3 albedo is calculated for each pixel, and if the albedo is < 3% th array is restored to the clear category.
C3AT	Channel 3 Albedo	The Ch3 albedo is calculated for each pixel, and if the albedo is > 6% th array is flagged as mixed or cloudy.
FMFT	Four Minus Five Test	A threshold, which is a function of Ch4, is calculated. If the difference between Ch4 and Ch5 is > the threshold, the array is flagged as mixed o cloudy.
RGCT	Reflectance Gross Cloud	If Ch1 reflectance is > 44%, the array is flagged as mixed (1-3 pixels) or cloudy (all 4 pixels flagged).
RRCT	Reflectance Ratio Cloud Test	For each pixel in the array, the ratio of Ch2 to Ch1 is calculated. If the ratio is between 0.9 and 1.1 the array is flagged as mixed or cloudy.
RUT	Reflectance Uniformity Test	If the difference between the minimum and maximum Ch1 of the array i > 9%, the array is flagged as mixed.
TGCR	Thermal Gross Cloud Restoral	If the Ch4 is >293 K, the pixels are restored to clear.
TGCT	Thermal Gross Cloud Test	If Ch4 is < 249 K, the array is flagged as mixed or cloudy.
TUR	Thermal Uniformity Restoral	If the maximum and minimum Ch4 difference is <1°K, the pixels are restored to the clear category.
TUT	Thermal Uniformity Test	If the difference of the minimum and maximum Ch4 btemp in the 2x2 array is > 3 K, the array is flagged as mixed.

The determinate test listed in Table 2-3 indicates the data condition that determined that pixels are cloudy and mixed. For clear values of 23–30, one or more of the tests determined that the pixels were cloudy or mixed; however, a restoral test later determined these to be clear.

For people familiar with the NOAA implementation of CLAVR, please note that the numeric values found in the Pathfinder data do not all match those used by NOAA. A mapping of NOAA CLAVR values in the Pathfinder data to the NOAA values is also provided in Table 2-3. The numeric values were reorganized to simplify value lookups for users who are interested in simple discrimination between clear (values 22–30), mixed (values 12–21) and cloudy (values 1–11) pixels.

The CLAVR flags are calculated using top of atmosphere reflectances, which are different from the surface reflectances stored in the data set data layers as described in Appendix B.

Table 2-3
CLAVR Flag Values

Flag Value	Condition	Determinate Test	NOAA Value	Flag Value	Condition	Determinate Test	NOAA Value
0	no decision		0	18	mixed	TGCT	24
1	cloudy	RGCT	1	19	mixed	FMFT	25
4	cloudy	RRCT	4	20	mixed	TUT	26
5	cloudy	C3AT	5	21	mixed	TGCT	27
6	cloudy	FMFT	6	22	clear		12
8	cloudy	TGCT	8	23	clear	TGCR	13
9	cloudy	FMFT	9	24	clear	TGCR	14
11	cloudy	TGCT	11	25	clear	TGCR	15
12	mixed	RGCT	17	26	clear	TUR	16
13	mixed	RUT	19	27	clear	C3AR	28
14	mixed	RRCT	20	28	clear	C3AR	29
15	mixed	C3AT	21	29	clear	TGCR	30
16	mixed	FMFT	22	30	clear	C3AR	31
17	mixed	TUT	23				

Multiple occurrences of a test are due to the test, identification, and restoral logic used by the CLAVR algorithm.

- **Quality Control Flags** (QC flag)— The QC flag is based on error or warning conditions that are identified during processing. The basic QC flags are shown in Table 2-4.

  These conditions are encoded in an additive fashion so that any specific condition or combination of conditions can be used and easily interpreted. For example, a QC flag with a value of **5** indicates that both the "filled data gap" bit **4** and the "Channel 1, 2 processing nonstandard" bit **1** were set.

Table 2-4
Quality Control Flag Values

Value	Condition	Condition
0	Normal	
1	Channel 1, 2 processing nonstandard	Ozone values were not available so climatology was used.
2	Channel 3, 4, 5 processing nonstandard	Calibration coefficients were unavailable.
4	Filled data gap	A data gap resulting from the forward transform used in binning has been filled with an adjacent pixel.
8	Range Check failure	Calculated values were outside the range of values.
16	NOAA QC flag set	See the "NOAA Polar Orbiter Data User's Guide" (Kidwel 1991).

Cumulative quality control flags are shown in Table 2-5. It should be noted that these conditions do not always indicate invalid values. Some QC flags, such as a range check failure, do mean that data in at least one layer are incorrect; however, some error conditions are informational and do not necessarily mean the data are not usable.

Table 2-5
Cumulative Quality Control Values

Flag Value	Conditions	Flag Value	Conditions	Flag Value	Conditions	Flag Value	Conditions
3	A, B	12	D, C	20	E, C	27	E, D, B, A
5	C, A	13	D, C, A	21	E, C, A	28	E, D, C
6	C, B	14	D, C, B	22	E, C, B	29	E, D, C, A
7	D, A	15	D, C, B, A	23	E, C, A, B	30	E, D, C, B
9	D, A	17	E, A	24	E, D	31	ALL
10	D, B	18	E, B	25	E, D, A		
11	D, A, B	19	E, B, A	26	E, D, B		

A = Ch 1,2 nonstandard processing
B = Ch 3,4,5 nonstandard processing
C = Filled data gaps
D = Range Check — value out of bounds
E = NOAA QC flag set

- **Sensor Scan Angle** — Sensor scan angle is the angle between the sensor view vector and the downward pointing (to nadir) axis. Figure 2-1 shows the solar/satellite geometry. The sensor zenith angle, which is not stored as part of the data set, is the angle between the sensor view vector and the local vertical to the pixel. The sensor zenith angle is used in the calculation of the relative azimuth.

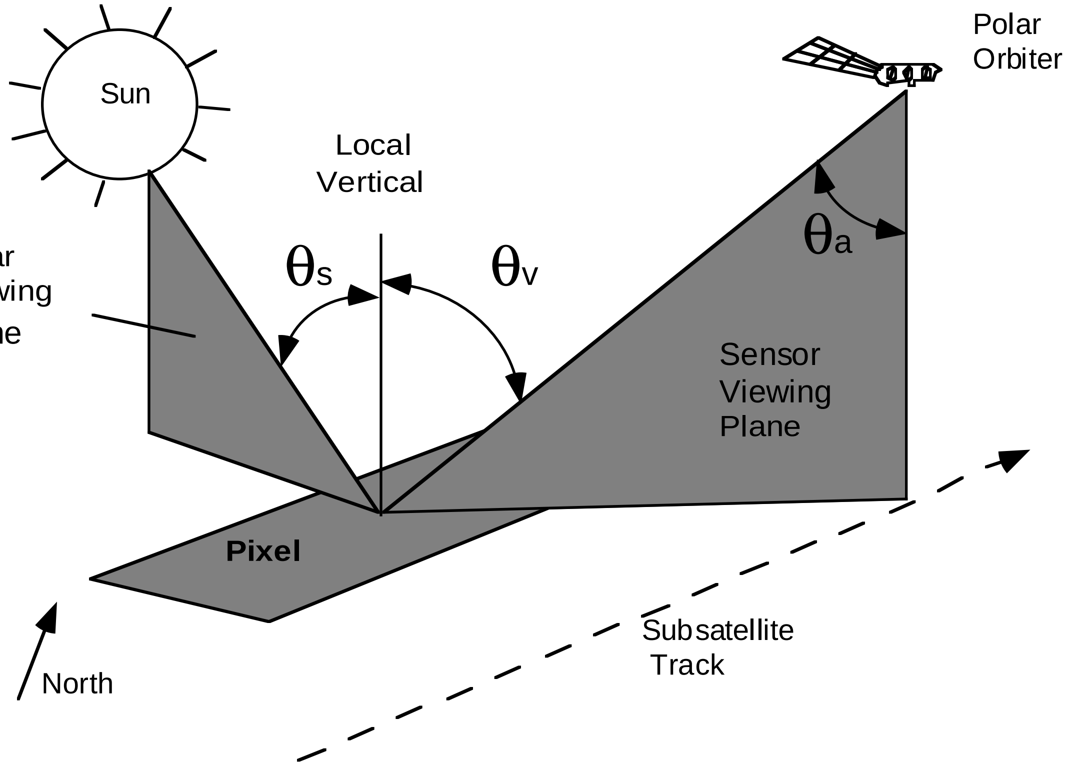
Scan angle is particularly important in that data with very large scan angles generally are greatly distorted and navigation of extremely off-nadir pixels may have large errors. Bidirectional effects are also more pronounced at extreme off-nadir pixels. In addition, at approximately 42 degrees of nadir the pixel width is greater than 8 km. For these reasons, in producing the Daily Data Set, pixels within 42 degrees of nadir are selected preferentially to pixels outside 42 degrees, and in the composite data set pixels outside 42 degrees are excluded.

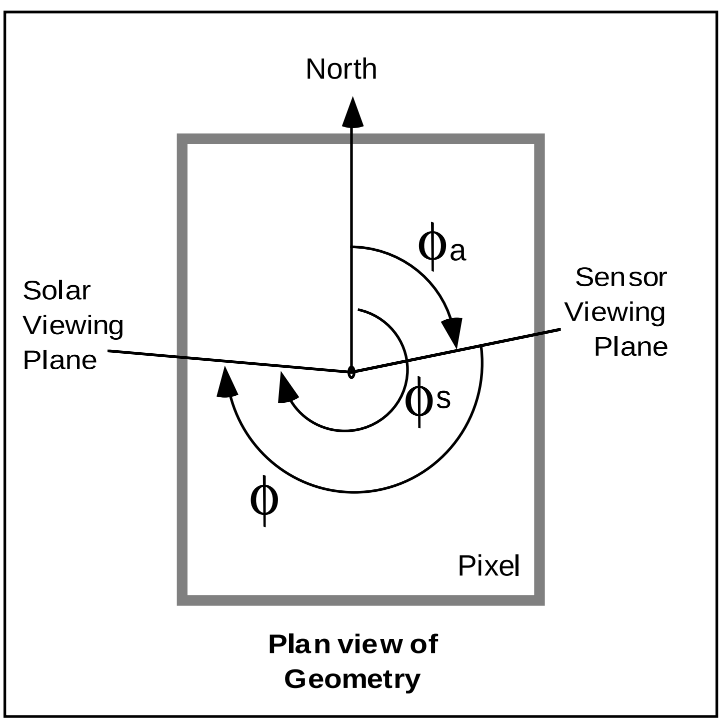
- **Solar Zenith Angle —** This is the angle between the solar vector and the local vertical to the pixel (see Figure 2-1).

  The precision of the AVHRR visible channels degrades rapidly in twilight areas. For this reason, all data with high solar zenith angles (i.e., where the Sun is close to the horizon) are discarded in processing. Specifically, all data with a solar zenith greater than 80 degrees are discarded. As a result, in winter, data from areas already in darkness at the satellite's overpass time (between about 14:40 – 17:30 local time) are discarded. However, because the solar zenith changes along a scan, only part of the data (the easternmost part) is discarded. Within the Daily Data Set, the resulting pattern is a "saw tooth" of missing data. This is most pronounced at the summer and winter solstices.

  This feature becomes extreme in the latter part of the NOAA–9 data (1987/88) when, due to orbital drift, the local time at satellite overpass is close to 6 PM and the satellite approaches darkness at relatively low latitudes. This also results in a strong contrast in Channel 1 and 2 brightness west to east across an orbit.

- **Relative Azimuth** — The relative azimuth angle is equal to the solar azimuth angle minus the sensor azimuth angle, where the azimuth angles are obtained by projecting the solar and sensor vectors onto the local tangent plane and measuring the angle from local north to the projected angle in the direction from north to east as shown in Figure 2-1.





θ_a = Satellite Scan Angle

θ_s = Solar Zenith Angle

θ_v = Sensor Zenith Angle

ϕ_a = Sensor Azimuth Angle

ϕ_s = Solar Azimuth Angle

ϕ = Relative Azimuth Angle

Figure 2-1. Solar and Scan Geometry

- **Channel 1 Reflectance** — This is the percent reflectance from the surface, derived from Channel 1 of the AVHRR (0.58 – 0.68 micrometer). The reflectance is corrected for Rayleigh scattering and ozone absorption.
- **Channel 2 Reflectance** — This is the percent reflectance from the surface, derived from Channel 2 of the AVHRR (0.73 – 1.10 micrometers). The reflectance is corrected for Rayleigh scattering and ozone absorption.
- **Channel 3 Brightness Temperature** — This is the top of the atmosphere brightness temperature, in Kelvin, derived from Channel 3 of the AVHRR (3.55 – 3.93 micrometers).
- **Channel 4 Brightness Temperature** — This is the top of the atmosphere brightness temperature, in Kelvin, derived from Channel 4 of the AVHRR (10.3 – 11.3 micrometers).
- **Channel 5 Brightness Temperature** — This is the top of the atmosphere brightness temperature, in Kelvin, derived from Channel 5 of the AVHRR (11.5 – 12.5 micrometers).
- **Date and Hour of Observation** — This field contains the date and hour of observation. The date is given as the day of year and the hour is given as GMT.

2.1.2 Additional Data Notes

Additional, important detailed information is provided in Appendix C, "Product Notes." Among the features of the daily data that may influence data analysis and interpretation, and which are described in these notes, are

- areas of missing data
- areas of bad data
- the contrast in the area of orbital overlap
- how data are filled in the area of orbital overlap
- CLAVR and QC flag details
- land/sea mask errors.

2.2 Composite Data Set

The Composite Data Set contains the same data layers as the Daily Data Set (Table 2-1, Section 2.1). Note that similar to the Daily data, as discussed in Section 2.1.1, the Composite data have 10-bit accuracy and the trailing 2 or 3 decimal places in output values do not represent increased data accuracy. The composite is generated by comparing the NDVI values for each 8 km bin from 10 consecutive Daily Data Sets. Because data at the edge of a scan may contain distortion and bidirectional effect biases, only data within 42 degrees of nadir are used in the composite. The pixel with the highest NDVI for the 10 days is chosen as the date for inclusion in the composite, and all 12 data layers are updated with data from that date. This compositing process is effective for removing most of the clouds and atmospheric contaminants, thus providing as close to a cloud free field in each of the data layers as is possible (Holben, 1986).

There are three composites per month. The first composite of each month is for days 1 to 10, the second composite is for days 11 to 20, and the third composite is for the remaining days. This convention was chosen so that these data could be used with many climatologies and meteorological data that are provided in monthly averages.

Some of the features described in the Daily Data Set are not present or are not as pronounced in the Composite Data Set. The contrast in the areas of orbital overlap is no longer present (because pixels outside 42 degrees are not used). In generating the 10-day composite, pixels flagged out of range or with NOAA QC flags (in the QC layer) are not included. This helps but does not totally eliminate the selection of data with abnormally high NDVI's resulting from bad calibration. However, there are frequent cases of good Channel 1 and 2 data where the thermal data are missing or incorrect due to a lack of calibration coefficients. These areas of bad data should be noted in the QC comments.

The user is again referred to Appendix C, "Product Notes," for additional details.

2.3 Climate Data Set

The Climate Data Set is a global 1-degree x 1-degree NDVI field. This is derived from the Composite Data Set by calculating a mean, cloud free Channel 1 reflectance and a mean cloud free Channel 2 reflectance for each 1-degree area in which 50% or more of the 8 km pixels are identified as land pixels. Only those pixels with the CLAVR flags of clear (i.e., 22 – 30) are used in the processing.

2.4 Browse Images

After the daily and composite products are generated, a reduced resolution browse product is generated for users to look at before ordering data.

Channel 2 reflectance and Channel 4 brightness temperature layers of the Daily Data Set are subsampled every 8th pixel and every 8th line. The data are rescaled to an 8-bit gray scale by removing 2% of the highest values and lowest values in each channel and performing a linear stretch on the data. For this reason, the daily browse data cannot be intercompared or scaled.

NDVI and Channel 4 layers of the Composite Data Set are subsampled every 8th pixel and every 8th line. Data in the Channel 4 brightness temperature layer are rescaled to 8 bit, and data with brightness temperatures less than or equal to 273 K are flagged with a value of 4, and data with values greater than or equal to 315 K are assigned a value of 254. A linear stretch from 273 to 315 degrees is applied (covering a gray scale range from 5 to 253).

For all browse data, value 1 indicates ocean data, 2 indicates the interrupted space in the equal area projection, and 3 indicates missing data over land.

2.5 Ancillary Data File

The Ancillary Data File is essentially a tool to aid in interpreting the daily and composite data. In addition to the full resolution ancillary data file, six predefined continental boundaries were used to subset the it into areas that cover Africa, Asia, Australia, Europe, North and South America. The layers of the ancillary data set are presented in Table 2-6. The elevation data are from NOAA's 5-minute Earth topography data set (ETOPO5)

(NGDC, 1993) and the land sea flag is from NOAA's operational land sea mask (Stowe et al., 1991).

Table 2-6
Ancillary Data Set Layers

Data Layer	Units	Range
1 latitude	degrees	–90 to 90
2 longitude	degrees	–180 to 180
3 global elevation/bathymetry	meters	–15000 to 10000
4 land/sea mask	n/a	

2.6 Calibration Data Set

Each file is a binary subset of the NOAA GAC 1B orbit and has three types of records. The first is the terabit memory (TBM) header copied from the NOAA GAC 1B orbit file. The second is the data set header record copied from the orbit file. Following that are the calibration records, one for each scan line contained in the original orbit file. The forms of the TBM and data set header records are as described in the "NOAA Polar Orbiter Data User's Guide" (Kidwell, 1991).

Specific format information can be found in Section 4.4, "Calibration Data Set ."

2.7 Map Projection

The daily and composite data are mapped to a global 8 km equal area grid using the Goode Interrupted Homolosine projection shown in Figure 2-2. This projection is designed to reduce distortion of the major land masses while maintaining equal area (Steinwand et al., 1992; Steinwand, 1994). The browse data are also in the Goode projection, however these are subsampled 8 by 8, rows by columns.

See also Section 4.2.3, "DAAC Software Tools," and Section 5, "Frequently Asked Questions," for additional information on coordinate conversions and reprojecting data from the Goode projection.

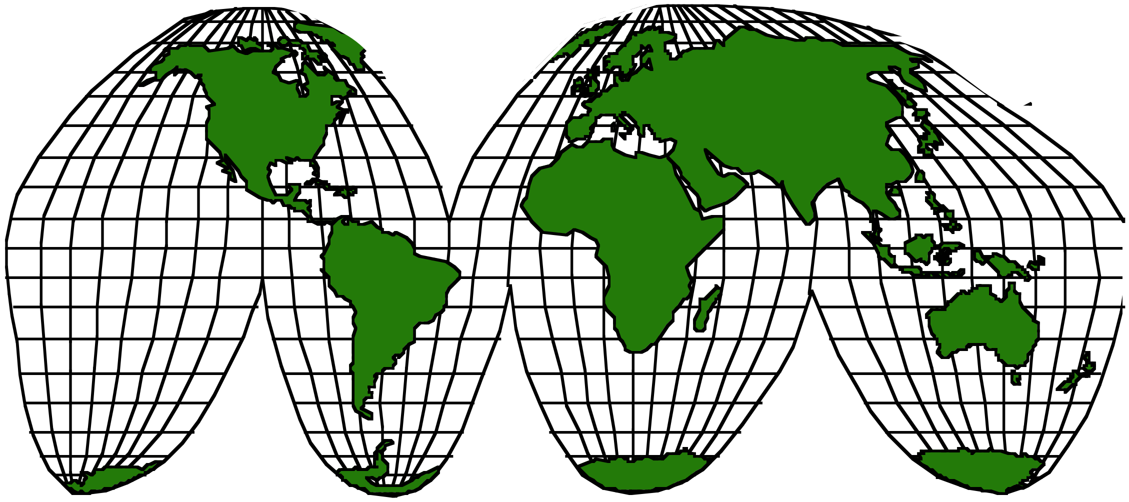


Figure 2-2. Goode Interrupted Homolosine Projection

2.8 File Naming Conventions

The daily files are named in the form month, day, year, using standard monthly abbreviations. For example, the image for April 1, 1988, has the file name

PAL_APR_01_1988.HDF

If the file is a composite, the file name is a hybrid of the first and last days in the composite. For example, the composite for the first 10 days of April 1988 would be named

PAL_APR_01-10_1988.HDF

In all file names, if a hyphen is found between two dates, the image is a composite.

The ancillary file name is

PAL_ANCILLARY_DATA.HDF

and the data set is invariant over time.

The browse file name is similar to the daily and 10-day files with the addition of "browse" in the name, e.g.,

PAL_BROWSE_MAR_20_1988.HDF

The climate data set follows the naming convention of the composite files; however, the term "climate" is inserted into the data set, e.g.,

PAL_CLIMATE_JAN_01-10_1988.HDF

In all files, a .Z appended to the file name indicates that the file has been compressed with the standard UNIX **compress** command. PAL is an acronym for Pathfinder AVHRR Land.

3. Archive and Data Access

The Goddard DAAC provides a variety of data support services for the Pathfinder and other data sets including

- Online data search and order
- FTP distribution of limited volume data sets
- Tape distribution of data sets
- Tools for reformatting and subsetting Pathfinder data sets
- User Services Office phone, fax, and email support
- Browse utilities that can run locally on your workstation and order data from the DAAC archive.

This section provides information on the available services and how to access them.

SECTION 3 CONTENTS

3.1 Goddard DAAC Services

The Goddard DAAC provides online search and ordering for the Pathfinder data and provides telephone, fax, and electronic mail consultation through the User Services Office (USO). Users can request data for network delivery via File Transfer Protocol (FTP) or on magnetic tape for surface mail delivery. Accessing the online system is easily performed using the Internet.

Access to Pathfinder data by FTP must be performed on line using the Goddard DAAC user interface (i.e., data are not available through anonymous FTP). Within the interface a user may select data files and documentation to be staged for FTP transfer. The daily and composite Pathfinder files are only available via FTP when compressed.

Since the DAAC systems and services operate in a "continuous improvement" mode, additional services and functionality not described in this document may be added. The DAAC USO can provide the most up-to-date documentation and information about DAAC services and systems.

3.1.1 DAAC Online System Access and Data Ordering

The DAAC system may be accessed using **telnet** and standard Transmission Control Protocol/Internet Protocol (TCP/IP). To connect to the system via **telnet**, type

telnet daac.gsfc.nasa.gov

If the message HOST UNKNOWN appears, type

telnet 192.107.190.139

to use the direct TCP/IP address.

The message

CONNECTED TO DAAC.GSFC.NASA.GOV

will be displayed to indicate that you are connected to the system.

After you have established **telnet** access to the system, you will be prompted to log into the system.

USER: daacims

PASSWORD: gsfcdaac

A series of informational messages will be displayed followed by the initial data system screen. You will be asked to enter your name. If this is your first online session, you will be asked to supply User Profile information. This information, such as your mailing address or fax number, will be stored and used for mailing data orders and other correspondence. After completing the User Profile, a tutorial will be displayed.

If you are unfamiliar with how to navigate within the online system, review the information in the tutorial. Moving through the online system requires use of the Escape key. To bypass the tutorial, enter

Esc Esc

The next time you log on, the main search screen will be displayed immediately after entering your name and confirming your contact information. Online help is available for all screens.

The Main Search Screen allows users to enter information such as data set type, dates of interest, locations of interest, etc. The data files retrieved as a result of these searches can be marked for ordering. If this is your first order of Pathfinder data, the Ancillary Data File will automatically be sent. If you need to order the Ancillary Data File specifically, select the Documentation/Software Ordering Screen from the Options button in the menu bar at the top of the screen.

3.1.2 Data Orders

If you submit a data request, you will be given an order identification number, which you may want to record in case you need to cancel or inquire about the order in the future.

If you submitted a request for data via FTP distribution, you will receive an electronic mail message indicating when your files are ready to be transferred. At that time you must log in again to the online system and you will be prompted to download the files. The DAAC generally stages files within one day.

If you submitted a request for data distribution on tape, ideally you should receive the tape directly from the DAAC within 3 working days. However, some large volume orders may take longer to process, but such orders will certainly be filled within 30 days.

3.1.3 User Support

For additional information or assistance, the Goddard DAAC USO can be reached as follows:

Hours: 9AM–5PM, Monday–Friday
Voice Line: 301–286–3209
Fax: 301–286–1775
Email: daacuso@eosdata.gsfc.nasa.gov

Goddard DAAC User Services Office
Code 902.2
Global Change Data Center
NASA Goddard Space Flight Center
Greenbelt, Maryland 20771

3.1.4 System Cost and Authorization

There is no cost to use the online system or to order data from the Goddard DAAC.

3.1.5 Tape Format and Access

Data are available on 4 mm (DAT), high- or low-density 8 mm (Exabyte), and 6250 bpi tape. Tapes are created with the UNIX utilities **dd** and **tar** on a Silicon Graphics 440 system. The no swap device and a block size of 63.5 KB are used by default, but users may request a convenient block size. Tapes may be requested in **dd** or **tar** file format. By default the data are archived and distributed in compressed format; however, they are available in uncompressed format by special request. Each tape distributed by the Goddard DAAC is labeled with the names of the files it contains in the order they were written. Files are compressed using the standard UNIX **compress** command, indicated by a .Z appended to the data file name, and should be uncompressed using the UNIX **uncompress** command.

Because the daily and composite files are 228 MB each, the DAAC recommends that files be copied from tape to disk before use. The same type of utility that was used to create the tape, **dd** or **tar**, must be used to access the data. If compressed files are requested, the data are compressed before the **dd** or **tar** utility is executed.

Therefore, compressed data should be uncompressed after copying from tape to local disk.

Appendix F contains system-dependent information for use in reading DAAC tapes on SGI, Sun, and VAX computer systems.

3.1.6 README File

A README file accompanies every shipment of data. In general, a DAAC data set README file is a text file containing a brief description of the instrument and its mission objectives, a summary of the specific data products available at the DAAC, an overview of the version of HDF used for the data products, and utilities developed by the DAAC for manipulating Pathfinder HDF files. It also contains caveats on the access, use, and quality of the data products. Finally, references to where additional information may be obtained are provided.

3.2 DAAC Software Tools

The Goddard DAAC has created software to enhance the use of the Pathfinder data set. This includes software for extracting geographic subsets of the data, and data to assist in reading the HDF format. These tools are designed for use on UNIX workstations only.

3.2.1 Description of Tools

Currently, there are three software packages available.

Pickacontinent — Continental subsetting software that allows the user to extract a predefined continental region from global, 8 km HDF files (dailies or composites) into a smaller HDF file. The coverages and dimensions of the output files are described in Section 4.

Hdftobin — Software that extracts one or all of the data layers from the daily or composite HDF file to a binary output data file. Input data can be a global daily or composite file, a continental HDF file (created from the pickacontinent routine), or a climate file. The output will be a binary data file without any header or the complementary metadata found in the original HDF file.

Goode — Provides the latitude and longitude coordinates for any specified bin in the daily and composite files (i.e., the 8 km data).

Getmeta — Extracts metadata from an uncompressed HDF file.

Newgetmeta — Extracts metadata from a compressed HDF file.

Pickacontinent and **hdftobin** use the HDF interface and libraries provided by NCSA. Specifically, the **hdfls** routine is used and must be installed on your workstation. For information on how to acquire and load the HDF software see Appendix E, "NCSA Hierarchical Data Format (HDF) Basics."

3.2.2 Accessing DAAC Pathfinder Tools

These DAAC Pathfinder software tools are available via FTP from the Goddard DAAC online system by request via the Document/Software List screen. It is distributed as a compressed **tar** file which contains the software as well as installation and use documentation. To receive the data, log into the DAAC online system as described in Section 3.1.1 and request the tools from the Document/Software List screen.

All of the DAAC Pathfinder tools are written in C. The source code, UNIX help files (**man** pages), execution scripts, and binary executable files are available in this directory. The binary files are for Silicon Graphics workstations only, and these tools have only been tested on workstations with a 32-bit word size. The tools (including **man** pages) will take less than 1 MB of disk space.

These tools will be available for distribution via the IMS in the Documentation/Software List screen.

3.2.3 DAAC Browse Tools

The Goddard DAAC Browser provides a way for the user to view browse images on the user's workstation and to order the daily and composite data set files associated with the browse image. Release 1.0 of the Browser is currently available with support for the current SGI, HP, and Sun operating systems (IRIX, HP-UX, and SunOS, respectively).

Browser 1.0 is available via FTP from the Goddard DAAC online system in the Document/Software List screen under Options on the menu bar. It is distributed as a compressed **tar** file that contains the software and installation and user documentation. To receive the Browser, log into the DAAC online system as described in Section 3.1.1 and request the software from the Document/Software List screen.

The software and sample files require at least 25 MB of disk space for installation. To run the Browser, your X Terminal or workstation should have at least 3.5 MB of memory.

To unpack the Browser, copy the compressed **tar** file from the DAAC online system to the destination directory on your workstation in which you want to install the Browser. The X11 and OSF/Motif libraries are required to build the Browser.

Type

zcat browse.tar.Z | tar xvf -

to unpack the software and related files. The "FILES" file is a list of all files that are included in the distribution. Installation instructions are contained in the README file.

Additional information on the Browser can be obtained through the DAAC User Services Office.

4. Reading the Data Sets

After obtaining Pathfinder data, you will need to understand how to read and interpret it. This section provides details on how to read the HDF-formatted data and how to scale them to derive geophysical values.

SECTION 4 CONTENTS

4.1 Getting and Using HDF

All Pathfinder data products (except for the calibration extracts) are in the Hierarchical Data Format (HDF) from the National Center for Super Computing Applications (NCSA). This is the standard data format for the Earth Observing System Data and Information System (EOSDIS) Version 0. Appendix E contains further information on HDF and how to obtain it.

4.1.1 Pathfinder Format Specifics

Each data file contains several types of information including data set descriptive information (metadata) and the scientific data. The actual pixel data are stored in separate bands, referred to as numeric *data layers*. HDF allows data (scientific data and metadata) to be implemented in several ways including Scientific Data Sets (SDS) and Raster Image Sets (RIS). The SDS implementation has more flexibility in including metadata and allows data of a variety of word sizes (8- to 64-bit data).

Table 4-1 summarizes the HDF format, dimensions, quantization, and sizes of the Pathfinder data products.

Table 4-1
Product Structure Overview

Data Product	Number of Data Layers - Quantization	Dimensions Pixels x Lines (columns x rows)	HDF Format	Approximate Size Compressed /Uncompress
daily data	3-8 bit, 9-16 bit	5004 x 2168	SDS	35 MB / 228 MB
composite data	3-8 bit, 9-16 bit	5004 x 2168	SDS	35 MB / 228 MB
ancillary data file	1-8 bit, 3-16 bit	5004 x 2168	SDS	18 MB / 50 MB
climate data	1-8 bit	360 x 180	SDS	12 KB / 66 KB
browse data	2-8 bit	625 x 271	RIS	95 KB / 342 KB

4.1.2 Metadata and Tags

The files that are in the SDS structure contain several types of information in addition to the scientific data. First are file annotations that contain the metadata for each specific file including information on the data set, data product, file size, geographic flag, day or night flag, file version, begin date, end date, producer identifier, satellite, quality control comments, and quality control date. This information is what is seen in the Goddard DAAC's online data base. An additional file annotation contains the quality control comments. In addition, there are object annotations associated with each SDS within the file that briefly describe the contents and units of each data layer. A calibration tag is attached to each layer of the SDS, and it provides the gain and offset information one should use to convert the scaled values to geophysical units (see Section 4.3).

4.2 Options for Image Display and Analysis

The Pathfinder Daily and Composite Data Set data files are very large (228 MB), and this presents challenges for viewing and processing. HDF access and viewing tools provide some benefit for easier data use; however, these tools are maturing and at present only work with the smaller Pathfinder files. Users often find that some combination of DAAC-provided subsetting or reformatting tools used together with off-the-shelf software is required to use the Daily and Composite Data Set.

4.2.1 NCSA Visualization Tools

The National Center for Supercomputing Applications (NCSA) provides several software packages that can be used to view some HDF data sets. Collage and Image for the Macintosh and XCollage for UNIX workstations are such packages. At present Collage and Image can be used to view the Browse Images only; the Daily and Composite Data Sets have 12 layers of data with dimensions of 5004 x 2168, which is currently too large for these tools to view. Collage can be used to display the HDF files created by the continental extraction software tool available from the DAAC. The functionality of these tools is increasing, and later versions are expected to be able to view all of the Pathfinder data sets.

To get the Collage and Image software, log into the NCSA anonymous FTP server using the command

ftp ftp.ncsa.uiuc.edu

Log into the Collage directory for Collage or the Mac/Image directory for Image using the command

cd Collage or **cd Mac/Image**

and ftp the files to your workstation using the FTP **get** or **mget** commands.

Then follow the instructions for installation and unpacking in the README files.

For users who prefer to read the data directly using the HDF calling interfaces, information on HDF access, installation, and use can be found in Appendix E.

4.2.2 Other Image Processing Tools

In addition to the tools mentioned above in Section 4.2.1, some off-the-shelf packages (commercial and public domain) can read information in HDF format. Among the software packages known to read some HDF format data are IDL, EASI/PACE from PCI, Inc., and Spyglass. You are encouraged to ask your vendors directly if they support HDF. Again, due to the complexity and size of the Daily and Composite Data Sets, software that reads some HDF format files may not read all of the Pathfinder files, and samples of Pathfinder data can be provided to vendors for test purposes.

4.2.3 DAAC Software Tools

One option users may prefer for viewing and analyzing the daily and composite files is to use the DAAC tools to extract all or part of the Pathfinder data to create binary files that can be read by their off-the-shelf or custom software. Information on how to download the DAAC tools is provided in Section 3.2.2, "Accessing DAAC Pathfinder Tools."

To use the DAAC tools, install them following the instructions in the README file and make sure the tools are present in your directory path. These programs will not work unless the HDF libraries are installed on your machine and the **hdfls** executable file (which is automatically created upon installation of HDF) is in your directory path. Shell scripts (with the **.sh** extension) are provided to execute the programs.

Using DAAC Tools

The following are some instructions and examples for using the DAAC tools once they are installed on your system.

hdftobin.sh

Enter

hdftobin.sh <hdffilename>

The following list appears

1 NDVI

2 CLAVR Flags

3 Quality Control Flags

4 Scan Angle

5 Solar Zenith Angle

6 Relative Azimuth Angle

7 Channel 1 Reflectance

8 Channel 2 Reflectance

9 Channel 3 BB Temperature

10 Channel 4 BB Temperature

11 Channel 5 BB Temperature

12 Day of Year

and you may select a layer number, **a** for all parameters, or **q** to quit.

The selected binary product layer will be extracted from the HDF file and written out to the same directory as the input file name. The binary files will have the same name as the input HDF file with the parameter name and number appended; however, none of the file metadata and tags will be present.

pickacontinent.sh

The boundaries of the six, predefined continental boundaries and the resulting file dimensions are shown in Table 4-2.

Enter

pickacontinent.sh <hdffilename>

and the following list appears.

0) Quit

1) Africa

2) Asia

3) Australia

4) Europe

5) North America

6) South America

Choose a number and an HDF file of the continental subset containing all 12 product layers and the same metadata and tags as the original file will be created in the same directory as the input file. It will have the same name as the input file with the continental name appended. Note that Hdftobin.sh can also work on continental subset files produced with the pickacontinent.sh to produce binary layers.

Table 4-2
Boundaries and Dimensions for Continental Extract Files

Area	North Latitude	South Latitude	West Longitude	East Longitude	# Pixels (columns)	# Lines (rows)	File Size
Global	90.0	–90.0	–180.0	180.0	5004	2168	227.8 MB
Africa	38.5	–37.8	–20.0	61.3	1100	1060	24.4 MB
Asia	79.0	4.5	25.6	143.7	1390	950	27.7 MB
Australia	7.5	–48.0	93.5	179.5	1080	770	17.5 MB
Europe	76.7	23.4	–13.0	60.5	780	670	10.9 MB
North America	72.5	9.0	–165.0	–60.0	1090	820	18.7 MB
South America	13.2	–57.0	–83.0	–33.0	690	970	14.0 MB

The Ancillary Data file has been subsetted for convenience, using the same six, predefined continental boundaries, and are available through the DAAC online IMS, in the Documentation/Software List screen.

goode.c

The goode.c file contains source code that will transform latitude and longitude coordinates into Pathfinder image coordinates and back again. The **demo.c** routine is an example of how such transformation code can be written. Type

demo

to run the program, enter a latitude and longitude as directed, hit return, and the program will give the forward transformation into the Pathfinder 8 km Goode coordinates. This program will also give the inverse transformation from the Goode coordinates back into the center latitude and longitude of the Goode bin.

getmeta

This is a quick way to look at the metadata in the files after obtaining the data. It will retrieve the metadata from the uncompressed HDF file and print it on the screen. Type

getmeta PAL_APR_01_1988.HDF

Another way to get this information is to view the metadata on line in the IMS.

newgetmeta.c

This will retrieve metadata from the archive HDF file without your uncompressing the entire file. Type

zcat PAL_APR_01_1988.HDF.Z | newgetmeta

4.3 Scaling the Data

The Pathfinder data are processed as 32-bit floating point numbers to maintain maximum accuracy. In generating the output data, each layer is scaled to an appropriate 8-bit (unsigned) or 16-bit (unsigned) integer value corresponding to the ranges shown in Table 4-3. Consequently, to obtain geophysical values from the scaled data value, offsets must be subtracted from the scaled data value and the result multiplied by the gain. Complete information on scaling and bit representations is provided in Table 4-3.

The scaling information (offsets and gains) for the Daily, Composite, and Climate data sets are stored in the HDF calibration tag and are referred to as calibration coefficients. This should not be confused with the calibration of the raw satellite counts. Using this information, which is stored in the HDF calibration tag, software could be written to automatically scale the data.

4.3.1 Scaling Examples

A For a binary value in the Channel 1 reflectance layer if the 16 bit value is 21009, using gain and offset values in Table 4-3, the reflectance is [(21009 – 10)(.002)] or 41.988 percent.

B Given a binary value from the scan angle layer of 8500, the scan angle becomes [(8500 – 10481.98)(.0001)] or –0.198198 radians. Convert radians to degrees (multiply by 57.2957795) and the scan angle becomes –11.36 degrees.

4.3.2 Data Flags

In each product there are certain values that indicate pixels are over ocean or missing data. Also, for the data in the Goode Projection and the browse images, the map projection interruptions carry a unique flag. These flag values, which do not require scaling, are also shown in Table 4-3. These flag values are less than the binary value of the minimum geophysical value as shown in Table 4-3.

4.3.3 Browse Images

The browse data are intended to be used only for scene selection. Because the daily browse images are each rescaled uniquely they cannot be intercompared, nor can they be scaled back to reflectance. Each image has a unique scale applied to provide some contrast between cloud and noncloud areas.

4.4 Calibration Data Set

Each file of the data set is a binary subset of a NOAA 1B file. There are three types of records in each file. The first record is the TBM header copied from the NOAA 1B file. The second is the Data Set header record copied from the NOAA 1B file. Following those are the Calibration records; one for each scan line contained in the original NOAA 1B file.

The TBM header record is 122 bytes long. The Data Set header record is 6440 bytes long. Each Calibration record is 488 bytes long. The number of calibration records varies from file to file. There is an average of about 13,000 Calibration records per file.

The form of the TBM record is described in Table 2.2.2-1 of the "NOAA Polar Orbiter Data User's Guide" (PODUG) (Kidwell, 1991), and the Data Set header record is described in Table 2.0.2-1 of the PODUG.

There are a few notable issues with the Data Set header record. The variable size listed in the PODUG is chosen to be 6440 bytes to match the data record length of a Full Copy Packed GAC record.

Table 2.0.2-1 of the PODUG lists bytes 83 – 6440 as zero-fill. Some of the original NOAA 1B files do not fill the entire record with zeros. The program used to write the Calibration data files forces all these bytes to zero.

The 448 bytes of each Calibration record are the first 448 bytes listed in Table 3.1.2.1-1 of the PODUG, page 3-4. These are in the form of packed GAC data records with the GAC video data and trailing spare bytes deleted.

The file names for the Calibration Data Set extracted files follows the form YYDDDHHMMSS.sat_cal. For example the filename: 87002141301.no9_cal indicates that the orbit is from 1987, day of year 002 (January 02), with the orbit beginning time of 14:13:01, and the file is from NOAA–9.

5. Frequently Asked Questions

5.1 About Data

1. *What is the schedule for Pathfinder processing?*

   It is estimated that 5 years of data will have been processed by December 1994. The full historical record from 1981 through 1993 will be completed by the end of December 1995.

2. *Is there any charge for large amounts of Pathfinder data?*

   Currently, there is no charge for any quantity of data. The DAAC encourages orders via FTP, but will also fill orders on tape. If an extremely large data order is requested on tape, the recipient may be required to provide blank tapes.

3. *How does the Pathfinder data set relate to other AVHRR Vegetation Index Data Sets?*

   Higher resolution 1 km Local Area Coverage (LAC) and High Resolution Picture Transmission (HRPT) data have been used for a variety of terrestrial remote sensing applications, but global 1 km coverage only became available from April 1992. Other related data products are the Global Area Coverage (GAC) 1B, NOAA's Global Vegetation Index (GVI) product (Kidwell, 1990; Goward et al., 1993), Global Inventory Monitoring and Modeling (GIMMS) product, and the European Community's Joint Research Centre (JRC) Monitoring of Tropical Vegetation (MTV) product. These data sets are discussed in Report 20 of the International Geosphere Biosphere Programme (Townshend, 1992).

   The Pathfinder builds on the techniques used to produce these data sets. The significant differences between the Pathfinder data and these precursors are that the Pathfinder data provides many data layers, provides daily as well as composite data, and provides global, mapped data at a higher resolution.

4. What is the Pathfinder Program and what are the other Pathfinder data sets?

Recognizing the need for improved, long time series data sets for global change research, NOAA and NASA initiated the "Early-EOS Pathfinder Data Set Activity" (aka The Pathfinder Program) in 1990. Candidate data sets were identified based on available coverage (temporal and spatial) and importance of the data for global change research. Four data sets were identified for the initial Pathfinder processing effort.

- Advanced Very High Resolution Radiometer (AVHRR) data from the polar orbiting TIROS/NOAA satellites
- TIROS Operational Vertical Sounder (TOVS) data from the polar orbiting TIROS/NOAA satellites
- Data from the Geostationary Operational Environmental Satellite (GOES)
- Special Sensor Microwave / Imager (SSM/I) data from the Defense Meteorological Satellites Program (DMSP).

In addition, NASA, through it's Mission to Planet Earth Program, initiated Pathfinder projects for

- Scanning Multichannel Microwave Radiometer (SMMR) from the Nimbus–7 satellite
- Selected time periods of Landsat data for land cover classification and change detection of several periods and regions.

In some cases there are multiple processing sites for these Pathfinders.

The TOVS pathfinder data are also available for ordering at the Goddard DAAC both via the online IMS and through the User Services Office.

5. Where can I obtain AVHRR GAC data?

NOAA/National Environmental Satellite
Data and Information Service National
Climatic Data Center
Satellite Data Services Division
Room 100
Princeton Executive Square

5267 Allentown Road
Camp Springs, MD 20746

Phone: (301) 763-8402
Fax: (301) 763-8443

Internet: sdsdreq@pes.sdsd.ncdc.noaa.gov

6. What ancillary data are required to produce Pathfinder data and where can I acquire these data?

The input ancillary data are described in Table 5-1.

The input ancillary files (ephemeris, clock resets, metadata, ozone, and topolandsea) can be obtained from the Pathfinder Processing Team. Files containing a list of orbits to be processed and a process control file are generated by users at the time of the run. NOAA 1B data are available from NOAA/NESDIS; however, at present users would have to provide some simple reformatting of the 1B data or rewrite the Pathfinder software to ingest standard NOAA 1B format data.

Table 5-1
Ancillary Data Sets

Ancillary Data	Description
Topography data	Land surface elevation from the Earth Topographic 5-Minute Grid (ETOPO5) data set (NGDC, 1993)
Ozone data	Daily 1-degree ozone data from the Total Ozone Mapping Spectrometer (TOMS) (Bowman and Krueger, 1985)
Metadata ancillary files	ASCII files containing both metadata (units, layer names, etc.) and DAAC metadata (listing general product information)
Land/sea mask	The NOAA operational land sea mask (Stowe et al., 1991).
Ephemeris file	Files containing orbital elements for each satellite available from NORAD
Clock reset file	One ASCII file for the full processing period — contains correction information for the NOAA onboard clock

7. How can I obtain NOAA ephemeris data?

The most current orbital elements from the North American Air Defense Command (NORAD) two-line element sets file are carried on the Celestial Bulletin Board Service, 513–253–9767, and are updated several times a week. Documentation and tracking software are also available on this system. Element sets (updated weekly) and some documentation and software are also available via anonymous FTP from **archive.afit.af.mil** (129.92.1.66) in the directory **pub/space**.

8. Some orbits are missing in the data that I received from the DAAC. What is the status and extent of this problem?

Occasionally some orbits are missing from the GAC 1B data used to create the Pathfinder AVHRR data set. Consequently, the processed data also do not have data in these areas. The quality comments in the metadata (accessible on line in the DAAC IMS) include information about missing orbits. Refer to these or the browse files for an indication, and extent of the occurrence for each data file.

5.2 About Software

9. What is the Pathfinder AVHRR Land processing system?

The Pathfinder system produces 8 km, terrestrial global data from the GAC data from the NOAA–7, –9, and –11 AVHRRs. The system reads Pathfinder format GAC orbits (the Pathfinder format is essentially a reblocked GAC orbit with a special header), renavigates each pixel, calibrates the five channels, corrects Channels 1 and 2 for Rayleigh scattering and ozone absorption, converts Channels 1 and 2 to surface reflectance, and Channels 3, 4, and 5 to brightness temperature. The software then generates cloud flags and NDVI, and creates a 12-layer output HDF file with Goddard DAAC specified metadata. Software documentation (overview, data flow, data dictionary, and program description) is available from the Pathfinder Processing Team.

10. Can I obtain the Pathfinder processing software?

Yes. The Pathfinder processing software is part of the public domain and all source code is available to users. However, the software is "use at your own risk." NASA will not implement changes on request. At this time, NASA does not have the resources to provide support beyond collegial email consultations.

11. How large is the Pathfinder Processing source code?

The Pathfinder source code consists of 26,900 lines of ANSI C code. HDF and Common Data Format (CDF) software are also required to run the Pathfinder.

12. What hardware is required to run the Pathfinder processing software?

Pathfinder software has compiled and run successfully on the following UNIX workstations: SGI, HP, Sun, and IBM. However, recent operating system upgrades may or may not have introduced problems in compiling or running the code. The minimum recommended hardware configuration is

- 64 MB memory
- 1GB storage for input and ancillary files (assuming 14 NOAA GAC orbits will be processed on any given run)
- 600 MB storage for output files
- 300 MB for all sources, libraries, and configuration control files (optional).

13. Where can I get HDF and CDF?

HDF is available from NCSA via anonymous FTP at **ftp.ncsa.uiuc.edu**. The operational Pathfinder uses version 3.2r4. It is available in the **HDF/prev_releases** directory. The Pathfinder processing system has been tested successfully with more recent versions of this software.

CDF is available from NASA via anonymous FTP at **ncgl.gsfc.nasa.gov**. The operational Pathfinder uses version 2.2, which is available in the **pub/cdf/cdf22-dist** directory. The Pathfinder processing system has been tested successfully with more recent versions of this software.

Documentation and support for both format software packages are available from NCSA and NASA at the sites listed above. Neither the Pathfinder team nor the Goddard DAAC is responsible for maintaining or supporting either package.

14. How can I reproject the Pathfinder Data to other map projections?

The General Cartographic Transformation Package (GCTP) may be used to transform the Pathfinder data coordinates from the Goode projection coordinates to other desired projection coordinates. Also, the DAAC tool "goode.c," described in Section 4.2.3, may be used to transform latitude and longitude coordinates into Pathfinder data coordinates and vice versa. The transformed coordinates can then be used with image processing software to reproject the data in the desired projection.

A C programming language version (GCTPc), which contains several new projections that were not in earlier versions of GCTP, is now available from the EROS Data Center (EDC) via anonymous FTP. Communications concerning GCTPc may be directed to

gctpc@edcserver1.cr.usgs.gov

Periodic updates and corrections will be put in this directory for access by the science community. Descriptions of these updates will be added to the README file in the main directory.

The package has been tar'd and compressed under the directory name **gctpc**. To access it, retrieve the file **gctpc.tar.Z** by anonymous FTP to

edcftp.cr.usgs.gov

This file can be retrieved by following these steps.

ftp **cd pub/software/gctpc**

ftp **get gctpc.tar.Z**

Then on a UNIX system, type

% **uncompress gctpc.tar.Z**

% **tar -xvf gctpc.tar**

The gctpc.tar file is 40 MB in size. A GCTPc directory will be created and the files copied into it. The installation will require 80 MB of disk space. The gctpc.tar file is not automatically deleted.

In order to reproject the Pathfinder data you may need to know the projection parameters for the global data set. These are listed in Table 5-2.

__Table 5-2__
__Goode Interrupted Homolosine Global 8 km Projection Parameters__

Number Output Lines 2168
Number Output Pixels 5004
Upper Left Y Coordinate 8669500.0
Upper Left X Coordinate –20011500.0
Output Pixel Size 8000
Earth's Radius (R) = 6370887.0 meters

Central meridians for each of the 12 regions			False eastings for each of the 12 regions	
lon_center[0]	–1.74532925199	–100.0 degrees	feast[0]	R * –1.74532925199
lon_center[1]	–1.74532925199	–100.0 degrees	feast[1]	R * –1.74532925199
lon_center[2]	0.523598775598	30.0 degrees	feast[2]	R * 0.523598775598
lon_center[3]	0.523598775598	30.0 degrees	feast[3]	R * 0.523598775598
lon_center[4]	–2.79252680319	–160.0 degrees	feast[4]	R * –2.79252680319
lon_center[5]	–1.0471975512	–60.0 degrees	feast[5]	R * –1.0471975512
lon_center[6]	–2.79252680319	–160.0 degrees	feast[6]	R * –2.79252680319
lon–center[7]	–1.0471975512	–60.0 degrees	feast[7]	R * –1.0471975512
lon_center[8]	0.349065850399	20.0 degrees	feast[8]	R * 0.349065850399
lon_center[9]	2.44346095279	140.0 degrees	feast[9]	R * 2.44346095279
lon_center[10]	0.349065850399	20.0 degrees	feast[10]	R * 0.349065850399
lon_center[11]	2.44346095279	140.0 degrees	feast[11]	R * 2.44346095279

15. Whom can I ask for limited Pathfinder science and processing information?

Mary E. James
Code 902.3
Global Change Data Center
NASA Goddard Space Flight Center
Greenbelt, MD 20771

Phone: 301–286–2432

Fax: 301–286–1775

Internet: mary.james@gsfc.nasa.gov

Appendix A - Science Working Group Members

LAND SCIENCE WORKING GROUP

John Townshend, Chair
Room 1113, Lefrak Hall
Department of Geography
University of Maryland
College Park, MD 20742

Roni Avissar
Department of Meteorology & Oceanography
Rutgers University
Cook College—P.O. Box 231
New Brunswick, NJ 08903

Jeff Eidenshink
EROS Data Center
United States Geological Survey
Sioux Falls, SD 57198

Chris Brest
NASA/Goddard Institute for Space Studies
2880 Broadway
New York, NY 10025

Sam Goward
Room 1113, Lefrak Hall
Department of Geography
University of Maryland
College Park, MD 20742

Garik Gutman
NOAA/NESDIS/ERA12
World Weather Building
5200 Auth Road
Camp Springs, MD 20233

Chester Ropelewski
Climate Analysis Center W/NMC52
NOAA Science Center, Room 605
Washington, DC 20233

David Schimel
Univ. Corporation for Atmos. Research
P.O. Box 3000
Boulder, CO 80307

Dan Tarpley
NOAA/NESDIS
World Weather Building
5200 Auth Road
Camp Springs, MD 20233

Compton J. Tucker
NASA/Goddard Space Flight Center
Code 923
Greenbelt, MD 20771

CALIBRATION WORKING GROUP

C.R.N. Rao, Chair
NOAA/NESDIS
World Weather Building
5200 Auth Road
Camp Springs, MD 20233

Peter Abel
NASA/Goddard Space Flight Center
Code 920
Greenbelt, MD 20771

Chris Brest
NASA/Goddard Institute for Space Studies
2880 Broadway
New York, NY 10025

Robert Cess
Inst. for Terrestrial & Planetary Atmosphere
SUNY Stony Brook
Stony Brook, NY 11794–2300

Robert Evans
RSMAS/MPO
University of Miami
4600 Rickenbacker Causeway
Miami, FL 33149

Garik Gutman
NOAA/NESDIS/ERA12
World Weather Building
5200 Auth Road
Camp Springs, MD 20233

Yoram Kaufman
NASA/Goddard Space Flight Center
Code 913
Greenbelt, MD 20771

William Rossow
NASA/Goddard Inst. for Space Studies
2880 Broadway
New York, NY 10025

W. Frank Staylor
NASA/Langley Research Center
Mail Stop 420
Hampton, VA 23665–5225

Michael P. Weinreb
NOAA/NESDIS
World Weather Building
5200 Auth Road
Camp Springs, MD 20233

Agency Management and Implementation Teams

Martha Maiden	NASA HQ	NASA Program Manager
Arthur Booth	NOAA NESDIS	NOAA Program Manager
Mary James	NASA/GSFC	Pathfinder Project Manager
Peter Smith	NASA/GSFC	DAAC Pathfinder Manager
Patrick Agbu	HSTX	DAAC Science and User Support
Chuck Davis	HSTX	DAAC Pathfinder System Administrator
Bin Bin Ding	RDC	DAAC Pathfinder Software Lead
Andrew Griffin	HSTX	DAAC Software Lead
Susan Hart (*Jan 91 – Oct 93*)	SSAI	Algorithm Development Team
Mike Jones (*Dec 91 – Jun 94*)	GSC	Algorithm Development Team
Satya Kalluri	UMD CP	Pathfinder Validation Scientist
Shunlin Liang (*Aug 93 – Jun 94*)	UMD CP	Pathfinder Validation Scientist
Robert Mack (*Jan 91 – Dec 93*)	NASA/GSFC	Algorithm Development Team
Robert Rank	HSTX	DAAC Operations Team
Padmini Raviprakash	GSC	Algorithm Development Team
Dave Rosenfelder	SSAI	Operations Team
John Rosenfelder	SSAI	Operations Team (*Jan 91 – Dec 93*)
Angela Sigmund	GSC	Algorithm Development Team
Dave Wolf	UMD CP	Pathfinder Coordinator
Chris Wacker	HSTX	DAAC Operations Team

Appendix B - Processing Algorithm Description

The following information provides a detailed description of the processing methods used in generating the Pathfinder AVHRR Land Data Set. Most of this information is taken directly from the article "The Pathfinder AVHRR land data set: An improved coarse resolution data set for terrestrial monitoring." (James, M.E., and S.N.V. Kalluri. *International Journal of Remote Sensing,* 1994 – in press), with permission from the authors.

1. Processing Overview

Prior to beginning the production of the Pathfinder data set, the level 1b GAC data were acquired and transferred to readily accessible storage media. Twelve years of GAC data were stored on more than 30,000 magnetic tapes (9-track and 3480 cartridge) at various NOAA facilities. Groups at NASA/Goddard Space Flight Center (GSFC) and NOAA/National Oceanographic Data Center (NODC) transcribed the data from magnetic tape to 12-inch write once, read many (WORM) optical disks. More than 60,000 orbits were transferred to 420 6-GB optical disks, which are the working archive for Pathfinder AVHRR processing. An overview of the Pathfinder processing stream is shown in Figure B-1. The steps in the processing chain follow the consensus recommendations of the Pathfinder Land Science Working Group. The processing steps can be summarized as follows:

1. Ingest
GAC orbits are read from archival optical disks and staged to a magnetic disk. An orbit is opened and the header information is checked for correctness and the first 120 scans are read. Scan lines from the descending portion of orbits are skipped in the processing. The channel data and GAC ancillary data (e.g., gains and offsets) are unpacked and stored in memory for processing. Ancillary data such as surface elevation, ozone, and land/sea mask are appended to the data.

2. Navigation
Because of the volume of data for the entire Pathfinder period, interactive navigation was deemed impractical, and implementation of advanced autocorrelation methods added significant processing time when navigating global data. However, observed errors of several GAC pixels are commonly found in the NOAA-provided Earth location points in the GAC 1b data set. Therefore, the decision was made to navigate the data using an orbital model and updated ephemerides. A precision navigation system, available from the University of Colorado, provides navigation accuracy to within one GAC pixel (Baldwin and Emery, 1993), and was adapted for the Pathfinder processing. A file containing calculated time offsets for the onboard clock is used to update the times for

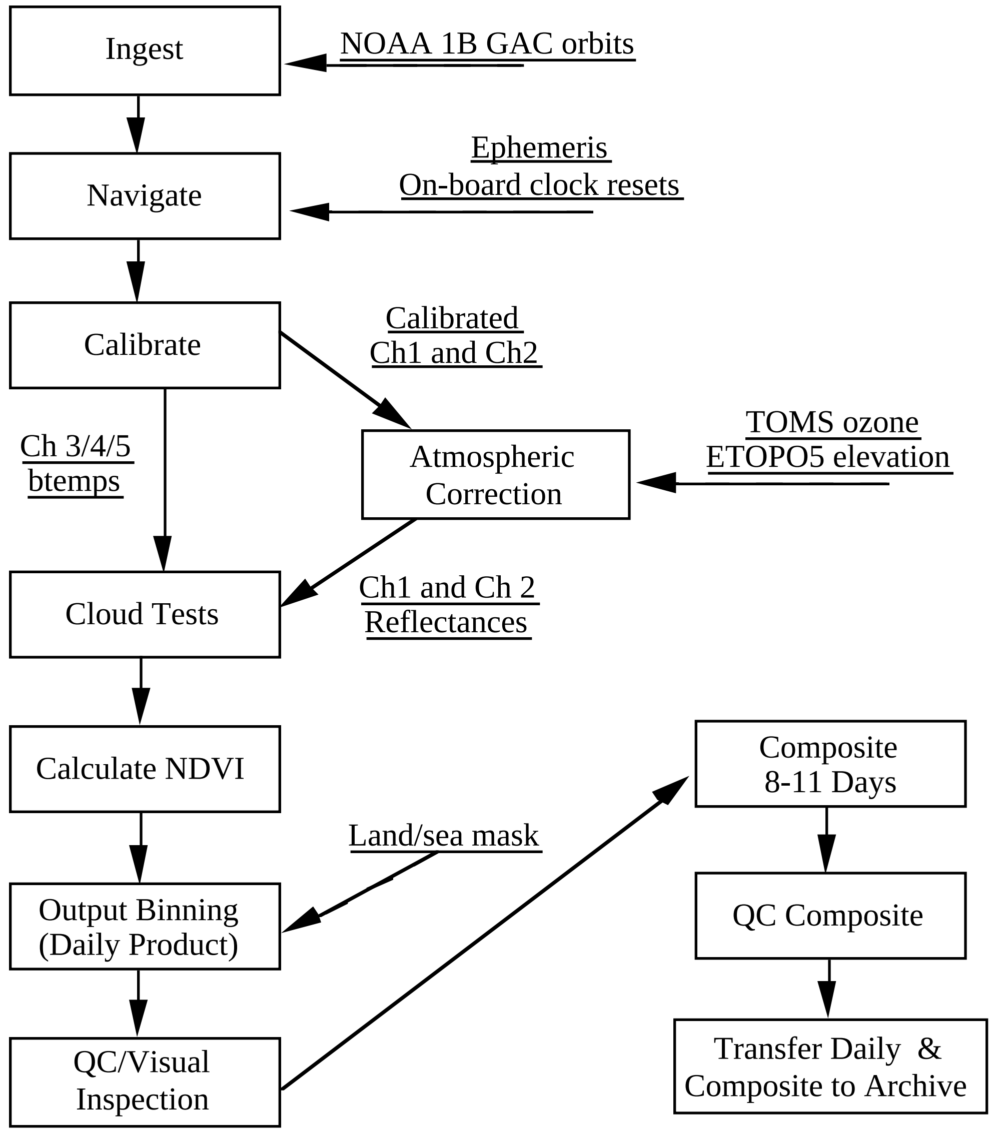


Figure B-1. Overview of Pathfinder Processing

each scan line, and revised Earth locations and solar/scan geometry are calculated for each pixel. In the Pathfinder processing chain, the Colorado navigation system is used to propagate the ephemeris data, and a closed-form geolocation algorithm developed at GSFC is used to geolocate the scan data (Patt and Gregg, 1994).

The Pathfinder processing team routinely checks the differences between the NOAA-provided tie points and the calculated locations and has found that the difference is generally within 6 to 8 km in latitude and 0 to 4 km in longitude with the greatest differences at the edge of scan.

3. Calibration (see Section 2)

Channels 1 and 2 are calibrated to produce at-satellite radiances using a time dependent correction that accounts for sensor degradation and intercalibrates among the satellites. Channels 4 and 5 are calibrated using a nonlinear function based on the internal calibration targets, baseplate temperatures, instrument dependent response curves, and NOAA-provided gains and offsets. Channel 3 is calibrated using the gains and offsets in the GAC data record. The thermal channels are then converted to equivalent brightness temperatures using a lookup table based on the inverse Planck function convolved with the instrument response.

4. Atmospheric Corrections (see Section 3)

A correction for Rayleigh scattering is applied. This correction takes into consideration ozone absorption (using ozone from the Nimbus–7 Total Ozone Mapping Spectrometer (TOMS)) and land surface elevation (from the ETOPO5 data set). Corrections including those for the eccentricity of the Earth's orbit and solar illumination are applied, and visible radiances are converted to percent reflectance.

5. Cloud Flagging

A cloud layer is calculated for and appended to each scan line using the Clouds from AVHRR (CLAVR) algorithm, which is being implemented by NOAA for operational processing (Stowe et al., 1991). This algorithm uses all five channels (top of atmosphere reflectances from Channels 1 and 2 and brightness temperatures from Channels 3, 4, and 5) to perform several tests using thresholds derived from sample data over a variety of surface types including deserts and ice fields. All tests are performed on a 2 x 2 array of pixels, and if all four pixels in an array pass all the cloud tests, each pixel in the array is flagged "cloud free." Otherwise, if 1 to 3 pixels fail cloud tests or if the array is highly variable (i.e., has strong contrast within the array) the array is called "mixed." The pixels are flagged "cloudy" if all 4 pixels fail the cloud tests. Some of the tests are used to restore pixels falsely identified as cloudy. The CLAVR algorithm is very conservative in identifying cloud free pixels, and pixels may be flagged as "cloudy" or "mixed" as a result of surface heterogeneity in areas of extreme elevation changes or along river courses.

6. Calculation of the NDVI

The NDVI is calculated using the difference ratio for the surface reflectances from Channels 1 and 2. Reflectances are used, rather than radiances or simple counts, to calculate the NDVI following Goward et al. (1991).

7. Binning and mapping (see Section 4)

Each GAC pixel is binned into one of the 8 km pixels of the output product. The bin is determined based on a forward, nearest neighbor mapping and the pixel selected as representative for that bin is the pixel with the greatest NDVI value for those pixels within 42 degrees of nadir. In the Daily Data Set the greatest NDVI value pixels outside of 42 degrees are used when there are no pixels within plus or minus 42 degrees. Forty-two degrees is used for a cutoff because at this point the GAC pixels are close to 8 km in width. Data for solar zenith angles greater than 80 degrees are not used in the daily or composite data because erroneous NDVI values are often found along the terminator (Holben, 1986) and the precision of the measured reflectances in twilight areas is very low (Goward et al., 1991). All data flagged as ocean are removed.

Steps 1-7 are repeated for consecutive groups of 120 scan lines until all scans and all orbits for one day are processed. Generally, one day's data are 14 orbits of about 13,000 scan lines each.

8. Quality Control

Throughout processing, automated checks are made for out of range values or missing values. Fatal error conditions result in excluding data from the final data set, and nonfatal error conditions are flagged per pixel. In addition, as each daily image is completed, the image is visually checked by an operator for obvious or gross errors. Any problems discovered are investigated and, if appropriate, corrected prior to generating the 10-day composite.

9. Compositing (see Section 5)

The Pathfinder data are composited for "quasi 10-day" periods. For each month, the first ten days are composited, the second 10 days are composited, and the remaining days of the month (8, 9, 10, or 11 days) are used to generate the final composite. The process of compositing several days of data to produce a cloud free data set is well understood. Higher

NDVI values tend to represent lower aerosol and cloud contamination (Holben, 1986) and the highest NDVI value is the composite method used in Pathfinder processing. It has been shown that simple highest NDVI compositing may result in scan angle biases due to bidirectional effects, so data with scan angles greater than 42 degrees are not used in generating the Pathfinder 10-day Composite Data Set, though these off nadir data are retained in the Daily Data Set.

2. Radiometry and Atmospheric Corrections

2.1 Calibration

A key goal of the Pathfinder processing is consistent calibration among satellites. For visible and near infrared channels, there is no in-flight calibration and there are marked differences in the stability and precision of the three AVHRRs on the three satellites being used in the Pathfinder processing. To enable intercalibration among visible and near-IR channels on the various satellites, several approaches have been used, including vicarious calibration techniques such as monitoring bright targets (e.g., deserts, areas of sunglint, high clouds) and limited aircraft underflights (Che and Price, 1992; Kaufman and Holben, 1993; Staylor, 1990; Smith et al., 1988). For thermal channels (Channels 3, 4, and 5), the presence of an onboard blackbody calibration target provides for continuous calibration, however much postlaunch work has been done to reanalyze the preflight calibration data and to model the nonlinear response of the detectors (Weinreb et al., 1990; Brown et al., 1985)

For the Pathfinder AVHRR processing, a calibration working group composed of AVHRR calibration experts from several disciplines was formed. The calibration procedures used in Pathfinder processing are based on recommendations of this group (Rao 1993a; Rao 1993b), and the members of the Calibration Working Group are listed in Appendix A.

Thermal and visible channel calibration is performed independently. As the data are read for processing, the satellite number, date of scan, baseplate temperatures and counts, internal calibration target counts, and calibration gains and offsets are retrieved from the GAC record and stored in memory for use in calibration. Channel 1 and 2 radiances are produced in units of $Wm^{-2}sr^{-1}um^{-1}$ and Channel 3, 4, and 5 radiances are calculated in units of $mWm^{-2}sr^{-1}cm$.

2.2 Visible Calibration

For the visible and near infrared channels, procedures outlined in Rao (1993b) are followed. The recommended offsets are tied to 1986 and 1988 U–2 aircraft underflights and gains are based on calibration trends determined using the southeastern Libyan desert as a time-invariant calibration target. Given the days since launch, t_d, calibrated radiance for each channel, L_i, is calculated as

$$L_i = (GAIN_i)(COUNT_i - OFFSET_i) \tag{1}$$

where the Gain is computed as

$$GAIN_i = A\exp(Bt_d) \tag{2}$$

where A and B, and $offset$ are listed in table B-1 (from Rao, 1993b, and Rao personal communication). These calibrated radiances are later input to the atmospheric correction.

Satellite	Channel	A	B	$offset$
NOAA–7	1	0.5753	1.01×10^{-4}	36.0
NOAA–7	2	0.3914	1.20×10^{-4}	37.0
NOAA–9	1	0.5406	1.66×10^{-4}	37.0
NOAA–9	2	0.3808	0.98×10^{-4}	39.6
NOAA–11	1	0.5496	0.33×10^{-4}	40.0
NOAA–11	2	0.3680	0.55×10^{-4}	40.0

Table B-1. Coefficients for Visible and Near Infrared Calibration

2.3 Thermal Calibration

Channels 4 and 5 are calibrated using coefficients and methods described in Rao (1993a). Before calculating brightness temperatures from Channels 4 and 5, the raw counts have to be converted to radiances, which in turn have to be corrected for the nonlinear behavior of the sensors.

First the raw counts (c) are converted to a "pseudo-linear" radiance, N''_{Lin}, using the linear equation

$$N''_{Lin} = G''c + I'' \tag{3}$$

where G'' is then gain given by

$$G'' = G' * \frac{(N_{ict} - N''_s)}{(N_{ict} - N'_s)} \tag{4}$$

where N_{ICT} is the radiance of the internal calibration target (ICT), G' is the gain specified in the level 1b data and N'_s and N''_s are correction terms provided to minimize the effect of the ICT temperature variations on the calculation following

Rao (1993a). N_{ICT} is calculated using a mean of the ICT temperatures over 50 consecutive scans. I'' is the intercept given by:

$$I'' = N_s'' - (G''/G')N_s' + (G''/G')I' \tag{5}$$

where I' is the intercept specified in the level 1b data.

Once N_{Lin}'' is calculated, the true radiance is calculated correcting for the nonlinear behavior of the sensor using the quadratic expression

$$N_{True} = A(N_{Lin}'') + B(N_{Lin}'')^2 + C \tag{6}$$

The coefficients A, B, and C for NOAA-9 have been computed by Rao (1993a) and are re-produced in Table B-2.

The calibrated fluxes are then converted to brightness temperatures. To save computation time, calibrated radiances are converted to temperatures using precomputed radiance to temperature conversion tables based on the response curves of each of the AVHRR thermal detectors.

No working group recommendations were made for the calibration of Channel 3, so the Pathfinder processing system uses the gains and offsets provided in the NOAA 1b data record and calibrates the radiances using the procedures described in the "NOAA Polar Orbiter Data User's Guide" (Kidwell, 1991).

The reader is referred to Brown et al. (1985), Weinreb et al. (1990), and Steyn-Ross et al. (1992) for further information on calibration of AVHRR thermal channels.

3. Atmospheric Correction

Considerable research is ongoing in the area of correcting AVHRR data for atmospheric effects. The effect of Rayleigh scattering is well understood, and significant advances are being made in understanding the effects water vapor and aerosol effects on AVHRR data (Vermote et al., 1990; Tanre et al., 1992). However, due to difficulties in finding ancillary water vapor and aerosol data sets and lack of experience in applying these corrections globally, the LSWG recommended that the Pathfinder only correct for Rayleigh scattering and ozone absorption in Channels 1 and 2.

Satellite	Channel	A	B	C	N''_s	N'_s	Time Period
NOAA–7	4	0.89783	.0004819	5.25	–5.16	–1.176	all
NOAA–7	5	0.93683	.0002425	3.93	–4.28	–1.346	all
NOAA–9	4	0.88643	.0006033	5.24	–5.53	–3.384	12/12/84 – 9/30/86
NOAA–9	4	"	"	"	"	0	9/30/86 – 11/7/86
NOAA–9	4	"	"	"	"	–3.384	11/7/86 – 10/16/87
NOAA–9	4	"	"	"	"	0	10/16/87 – present
NOAA–9	5	0.95311	.0002198	2.42	–3.06	–2.313	12/12/84 – 9/30/86
NOAA–9	5	"	"	"	"	0	9/30/86 – 11/7/86
NOAA–9	5	"	"	"	"	–2.313	11/7/86 – 10/16/87
NOAA–9	5	"	"	"	"	0	10/16/87 – present
NOAA–11	4	0.84120	.0008739	7.21	–8.05	0	all
NOAA–11	5	0.94598	.0002504	2.92	–3.51	0	all

Table B-2. Coefficients for Thermal Calibration

The first step in atmospheric correction is to calculate the reduced solar irradiance from the absorption of the ozone layer following Gordon et al. (1988). Daily ozone values from Nimbus–7 are used, and where the daily ozone values are missing, a mean value of 313 Dobson units is assumed and the QC flag for the pixel is updated to indicate that this value was used. The reduced solar irradiance after two passes through the ozone layer, F'_i, and one pass through the ozone layer, B'_i, are then calculated as

$$F'_i = F_i * \exp(-\tau_{oz} * [1.0 / \cos\theta_S + 1 / \cos\theta_v]) \tag{7}$$

$$B'_i = F_i * \exp(-\tau_{oz} * [1.0 / \cos\theta_S]) \tag{8}$$

where F_i is the top of the atmosphere solar irradiance corrected for day of year and eccentricity, τ_{oz} is the ozone optical thickness and θ_S and θ_v are the solar and sensor zenith angles. Surface reflectance, ρ_i, is then calculated as the ratio of atmospherically corrected ground leaving radiance (L_w) to corrected solar irradiance following

$$\rho_i = \frac{L_w}{B_{i'} \exp(-\tau_r (1 / \cos\theta_S))} * 100 \tag{9}$$

where

$$L_w = \frac{L_i^R}{(-\tau_r (1 / \cos\theta_v))} \tag{10}$$

and L_i^R is the calibrated flux minus the Rayleigh correction, which includes a small correction for pixel elevation following

$$[R(\theta_v, \theta_s, \phi) * F'_i] * \left[\frac{(1 - \exp(-\tau_{r'} / \cos\theta_v))}{(1 - \exp(-\tau_r / \cos\theta_v))} \right] \quad (11)$$

where τ_r is the atmospheric optical depth, and $\tau_{r'}$ is the optical depth corrected for the scale height of the observation based on elevation from the ETOPO5 data set.

4. Binning and Map Projection

There are many complex issues involving definition of the output grid and methods for selecting pixels to populate the bin in the output grid. The methods used in the Pathfinder process for mapping, binning, and navigation were selected based on LWSG recommendations and processing throughput limitations.

The Pathfinder daily and composite data are mapped to a Goode Interrupted Homolosine Projection at a resolution of 8 km (Steinwand, 1994). The output resolution of the Pathfinder data was chosen to maintain highest possible resolution using data up to 42 degrees off nadir. The map projection chosen was based on analyses performed by the USGS, in support of the IGBP Land Cover project, in which the Goode Interrupted Homolosine Projection was determined to provide the best global, equal area representation of the data while still allowing easy display and subsetting. Compatibility with the IGBP product map projection was also seen as a benefit to researchers.

Because the AVHRR scan data and the output 8 km grid are not affine surfaces and there are significant changes in scale, resolution, and orientation between the two data planes, the identification of output bins for each observed pixel must be based on the Earth location of the satellite pixel and the output bin. The method of determining into which output bin a satellite pixel is mapped is based on the "nearest neighbor, forward binning" approach. In this approach, the center latitude and longitude of each satellite pixel is used to determine into which output 8 km bin the satellite pixel maps. At nadir there may be as many as 6 pixels mapping to one output bin; however, the general case is 4 pixels per bin. In some cases around the equator no satellite pixel will be mapped to an output bin. This resultant Moiré pattern of null values has to be taken into consideration in image display and quantitative analysis.

In the case of the Pathfinder processing, forward binning was chosen primarily because of processing throughput considerations, due to the fact that no data gaps (holes) are present in data within 42 degrees of nadir, and, as Justice et al. (1989)

show, the degradation of the data introduced by the onboard GAC sampling are as significant as the effects of forward or inverse binning. The Pathfinder processing simply copies the pixel to the east of a data gap to fill small data gaps (i.e., gaps of 2 or fewer pixels) in the daily data, and a QC flag is set for all of those values that were copied to fill data gaps. To maintain physical values derived from the satellite, no data smoothing or interpolation is performed As a result, in areas of orbital overlap a checkered pattern is often formed due to cloud fronts moving in the 101 minutes between orbits or extreme bidirectional effects.

The remaining consideration in forward mapping is how to select which pixel values are retained in the output bin for the daily data. This occurs most frequently at high latitudes where there is daily orbital overlap, and at all latitudes near nadir where the input pixels are significantly smaller than the output bins. Following the reasoning that maximum NDVI pixels generally represent the pixels with the clearest atmosphere, the selection of the maximum NDVI pixel is being used in the Pathfinder processing.

5. Compositing

Once all daily data sets in a compositing period (for each month this is defined as day 1–10, 11–20, and 21–end of month) have been processed, an NDVI composite is formed. This is often referred to as a cloud free composite because the compositing procedure is designed to reduce the amount of clouds and dust in the data set. To do this, the NDVI and scan angle layer of all 10 (or so) images are read into memory. For each bin, the date with the highest NDVI, for those pixels within 42 degrees of nadir, is identified and all 12 layers of the output image are updated with the 12 channels from the selected date.

Appendix C - Product Notes

The following information about the Pathfinder AVHRR Land Data Set is intended to explain to Pathfinder data users some of the unique features of the Pathfinder data set. These notes contain very important information about the data that **all investigators should review before using**

Pathfinder data for research purposes. Information presented in these notes deals exclusively with the geophysical data; information specific to scaling, reading, subsetting, and remapping the data can be obtained from the Goddard DAAC either through the online system, the NOAA/NASA Pathfinder AVHRR Land Data Set User's Manual, or contacting the User Services Office.

This information has been prepared by the Pathfinder algorithm development team and is intended to supplement the user's manual.

Appendix C - Product Notes

The following information about the Pathfinder AVHRR Land Data Set is intended to explain to Pathfinder data users some of the unique features of the Pathfinder data set. These notes contain very important information about the data that **all investigators should review before using Pathfinder data for research purposes.** Information presented in these notes deals exclusively with the geophysical data; information specific to scaling, reading, subsetting, and remapping the data can be obtained from the Goddard DAAC either through the online system, the NOAA/NASA Pathfinder AVHRR Land Data Set User's Manual, or contacting the User Services Office.
This information has been prepared by the Pathfinder algorithm development team and is intended to supplement the user's manual.

A. Uncharted Territory

Prior to the Pathfinder data set, global, daily, mapped NDVI and associated channel data have been available to very few researchers. Rather than creating a highly refined, single channel NDVI product, the Pathfinder seeks to maximize available data to foster new data application research and new AVHRR data derivation and compositing methods. Having a daily data set introduces many data conditions which will be unfamiliar to users; orbital overlap, solar zenith cutoffs, missing data, etc., all have characteristic patterns that need to be understood. Similarly, the fact that in addition to NDVI, one has thermal, reflectance, cloud, and geometry information available for research presents many problems and opportunities. Data that at first glance appear "noisy" or "suspect" may appear in one or more data layers, but after reviewing this information, it is hoped that you will understand the causes and consequences of the data as found in the Pathfinder data sets.
When beginning an investigation with Pathfinder data, it is critical to spend a little time looking at the data you'll be using in your research. You should also review these notes and the "Data Set Details" (Section 2) in the NOAA/NASA Pathfinder AVHRR Land Data Set User's Manual. The following sections describe some of the known data conditions that may result in unexpected and undesired research results. Most of the data issues described in these notes are being investigated to determine the effect of possible modifications in future reprocessing.

B. Missing Data

1. **Missing scans and orbits.** If you are working with daily data, watch for missing orbits and scans. Many daily data sets contain data gaps ranging from a few scan lines to an orbit or more. In some cases, this is because orbits were lost in acquiring the Pathfinder 1b input data stream. In other cases, part of an orbit may be fine, but the rest of the orbit may not be processed because of

unrecoverable errors such as bad scan times (if the scan time is invalid it cannot be navigated). In our experience, we have found that about one out of 20 orbits contains some bad scan time or date information, and about 1 out of 12 days is missing an orbit. Users can determine if data are missing either by viewing the quality control comments (available on line at the DAAC and in the data file) or by viewing the browse data.

2. **Missing data "wedges."** When a single orbit is missing, the mapped result will have a wedge shape since as latitude increases the amount of overlap between orbits increases. As a result, north or south of about 55° latitude the area of the missing orbit(s) will be filled in with data from adjacent orbits.

3. **Seasonal "saw tooth" pattern.** The precision of the AVHRR visible channels degenerates rapidly in twilight areas. For this reason, all data with high solar zenith angles (i.e., where the Sun is setting) are discarded in processing. Specifically, all data with a solar zenith greater than 80° are discarded. As a result, in winter where areas are already in twilight at the satellites' overpass times (between about 14:40 – 17:30 local time) the data are discarded. However, because the solar zenith changes along a scan, only part of the data (the easternmost part) is discarded. The resulting pattern is a "saw tooth" of missing data. This is of course most pronounced at the summer and winter solstices.
This feature becomes extreme in the latter part of the NOAA–9 data (1987–88) and the NOAA–11 data (1993–94) when, due to orbital drift, the local time of equator crossing is close to 5 PM and the satellite approaches darkness at relatively low latitudes. This also results in a strong contrast west to east along orbit for the visible channels.

4. **"Filled" data gaps.** In mapping from satellite scan coordinates to the output 8 km mapped data, the center latitude and longitude of the scan pixel is used to determine the output pixel. Because the pixels at the edge of scan may be 6 km x 12 km or larger, there are some output bins to which no scan data map. To generate

an uninterrupted data field, data gaps of 1 or 2 pixels are "filled in" as a final production step. This is done by simply checking for missing data over land (data value of 0 at x, y) and filling in that value of the adjacent pixel to the east (x-1, y). If the pixel (x-1, y) is also missing, the next pixel to the east (x-2, y) is copied to both missing pixels. No value is copied to fill more than two pixels. In all cases, the QC layer bit is updated to indicate "filled data." In order to maintain physical values derived from the satellite no data smoothing or interpolation is performed — pixel values are simply replicated. This may enhance the "checkerboard" pattern described later in these notes. If this will affect your research, you may want to consider restricting your data to close to nadir observations (using information from the scan angle layer in the Pathfinder data set) or, if applicable, running a noise filter.

C. Quality Control Flags

Users should watch these flags as they may be indicators of major problems. For example, bad calibration information will often (but not always) be identified by a NOAA QC flag (numeric value 16), and pixels with "out of range" flags may be invalid. Since the flags do not indicate *which* data layer has a problem, it is best to either view the data you are using, or, at a minimum, read the qc_comments (which may be viewed in the DAAC online Information Management System or by the browse system, or by extracting metadata from the file) as the comments will generally indicate which layers have a large number of QC flags or bad values. If your analyses are sensitive to noise, you may choose to mask out all values with QC flags, particularly the NOAA QC and the DATA OUT OF RANGE flags.

SCALE the QC flags! You must subtract 1 from the data set value to determine the actual QC flag value (in 8 bit layers, it's easy to forget to do this!).

D. Navigation and Land/Sea Mask Errors

The land/sea mask used in Pathfinder processing is the NOAA "operational" land/sea mask, which was acquired by the Pathfinder team. This mask has nominal 6 km resolution, but due to the geometry of mapping from the mapped, 8 km bin center to the NOAA file, a general shift of one or two pixels is observed, particularly along northern coastlines. This shift is consistent throughout processing and therefore is not thought to be a navigation error. Exceptional shifts of 2 to 3 pixels are noted in the QC comments metadata; larger shifts are assumed to be navigational errors and are investigated and corrected. This shift is not likely to be detected in most continental or global scale analyses; however, users may wish to use the land/sea mask in the ancillary data to "grow" a larger mask if contaminated pixels might affect research.

Another apparent navigation error occurs at the edge of scan pixels. These pixels are far larger than the 8 km output bins, and any spacecraft attitude adjustments (e.g., jitter) will result in the most extreme misnavigation at the edge of the scan. Occasionally in the data you may observe large coastline shifts of 3 or so pixels (which is different from the systematic coastline shift described above), and in most cases these are at the edge of scan. In areas where there are sharp changes in surface type (e.g., along the Nile valley), the selection of the greenest pixel, combined with orbital overlap at the edge of scan, may result in "ghosting" where the surface edge appears twice (e.g., the Nile would appear twice with a one or two 8 km bin separation). Because these data are outside 42 degrees scan angle, this will have minimal effect on studies of vegetation index or surface characteristics using the Composite Data Set.

In addition, a few obvious errors exist (e.g., several ocean pixels near the Kamchatka peninsula are flagged as land) so again, check the area in which you are working. It was decided that no changes would be made to the land/sea flag in order to maintain a consistent data set throughout the Pathfinder period.

E. "Bad" Scans and Pixels

Occasionally, the counts in Channel 1 or 2 are incorrect. Frequently a NOAA QC flag is set a scan or two *after* the bad scans, but there is no indication of a problem for the specific bad pixel. When converted to reflectance, some of the bad pixels will be out of range (0 – 100%) and will be flagged as such, but basically the only way to determine if this condition exists is to inspect the data before analysis or read the qc_comments. These occasional bad values are the reason that pixels flagged as out of range are not used in compositing; however, many pixels that are within the valid range are still incorrect (e.g., seeing pixels with NDVI's of .8 in the middle of deserts or ice sheets). One artifact of compositing with the highest NDVI value is that bad pixels with unusually high values will be preferentially brought forward into the composites.

Schemes to detect bad counts and calibration are being developed; however, all currently proposed schemes involve regeneration of the 1b orbital data.

At times, calibration data for the thermal channels is either incorrect or unavailable. Rather than a calibration value, NOAA may provide a flag value, and because early in the processing (the data year 1987) this flag was not checked, bad temperature values (i.e., extremely low) were derived from these incorrectly calibrated data. These pixels generally contain a QC flag indicating that a NOAA QC bit was set. The valid range for brightness temperatures used in the Pathfinder (160 K – 340 K) was chosen based on a compromise of experiences of NOAA data users and the ISCCP processing team.

F. Edge of Scan "Checkered" Pattern

In the daily data, pixel selection is based on the greenest (i.e., highest) NDVI with preference being given to pixels within 42 degrees of nadir. When an orbit is mapped to the output 8 km bins, where there are no existing pixels the incoming pixels are binned to the output grid regardless of scan angle or reflectance. When the following orbit is mapped, data from the second orbit, which was imaged from a different viewing angle and was imaged 101 minutes later, are considered for binning. In the area of orbital overlap, the pixel from the second orbit will be binned if it has a larger NDVI value or if a data gap existed as a result of binning edge-of-scan pixels to the 8 km output. Frequently, a checkered pattern will result due to (a) different viewing geometry (bidirectional effects) or (b) cloud fronts moving into or out of an area in the 101 minutes between orbits. Since this area of orbital overlap is outside 42 degrees, these data are not included in the Composite Data Set.

G. CLAVR

The basic approach to the CLAVR layer should be *caveat emptor*. The CLAVR algorithm was chosen as an experimental cloud layer. This algorithm is very sensitive to surface inhomogeneity, and a single set of reflectance and thermal thresholds is used for the whole globe. The algorithm is presently not validated globally, so when using this cloud layer, users should read about the CLAVR algorithm and investigate which specific cloud flags will be important to your research.

CLAVR uses all 5 channels, so if any QC flags are set, it is wise to be suspicious of CLAVR — this is particularly true of data in which NOAA QC flags are set (as this is often an indication of bad thermal calibration). The Pathfinder implementation of CLAVR currently does not explicitly check for missing or fill values in tests using thermal channels and, as such, may calculate incorrect values where a data gap of a few scan lines is found. Also, the CLAVR values for 1987 data are slightly different from those in other data years due to an error in implementation that was not discovered because of the lack of available validation data sets.

When comparing a CLAVR value to the actual channel data of a pixel and surrounding pixels, it may not match the expected thresholds found in CLAVR. There are two reasons for this. First, the CLAVR value is calculated based on pixels in satellite scan resolution — when mapping to the output 8 km bins, some of the adjacent pixels may not be retained in the output data set. Also, CLAVR uses top of the atmosphere reflectances, while the Channel 1 and 2 reflectances given in the Pathfinder data set are atmospherically corrected surface reflectance.

Finally, you **must SCALE the CLAVR flags.** You must subtract 1 from the data set value to determine the actual CLAVR value.

H. Incorrect Nadir Values

A known problem exists with the Pathfinder Processing System, which results in the calculation of an incorrect scan angle and solar zenith angle at nadir for some pixels. This only occurs with the nadir pixels north of 55 degrees north latitude. Because these almost always have a scan angle of greater than 42 degrees these data are not included in the composites. Once this problem has been investigated and corrected, the Processing Changes report will be updated to identify the time period of data potentially affected.

I. Compositing

The Pathfinder data are composited for three "10-day" periods per month. The day with the highest NDVI value is selected for the Composite Data Set; however, because data at the edge of scan may contain distortion or bidirectional effect biases, only data within 42 degrees of nadir are used in the composite. It is important to note that in generating the 10-day composite, pixels flagged out of range (in the QC layer) are not included in the composite. This helps (but does not totally eliminate) the "compositing in" of data with abnormally high NDVIs resulting from bad Channel 1 and 2 calibration. However, there are frequent cases of good Channel 1 and 2 data where the thermal channels are missing or incorrect due to lack of calibration coefficients. If you are using the thermal channels of the composite, it is important to check the QC flags. **If there are residual areas of bad thermal data, and this might affect your results, you are advised to either generate a custom composite or use a different time period.**

J. Browse

Browse data are useful for a quick check of cloud cover and missing data in daily images, or cloud/snow contamination in the 10-day composite. The browse files are generated by subsampling every 8th pixel of every 8th line of two layers of the original data set. In the daily browse image, Channel 2 reflectances and Channel 4 brightness temperatures are used and for the 10-day composite browse the NDVI and Channel 4 are used.

For the daily browse images a histogram is calculated for each subsampled image, and 2% of the data points from each tail of the histogram are removed and an equalization stretch is performed on the rest of the values. The values at the tails of the histogram are then set to the remaining lowest/highest 8 bit gray level and the remaining data are scaled to 8 bit. Because of this, **gray levels in any one image do NOT correspond to the gray levels in any other image** and browse data should not be compared over a time series. In the 10-day composite browse, Channel 4 is scaled by setting all data below 273 K to a flag value, and all data above 315 K to another flag value, and then performing an equalization stretch on the remaining values. In composite Browse Images a color palette is included in which greener areas indicate more vegetation.

Some of the data conditions described in these notes may be observable in the browse. Among these are areas of missing data in the Daily Data Set, bad visible or thermal calibration in the daily files, bad thermal channels associated with good NDVI values in the composite file, and large areas of residual cloud and snow in the composites.

Appendix D - Processing Changes

A cardinal principle in Pathfinder product generation of long time series data is consistency in data processing. The Pathfinder AVHRR Land data set Processing Team spent significant effort in preoperational testing and validation, and instituted strict configuration and consistency controls. The policy to ensure consistency requires that no changes be made to the Pathfinder processing software that would affect the geophysical data and that any changes require approval from the Land Science Working Group prior to implementation.

Some abnormalities discovered in the processed data after evaluating the first year's data (1987) required changes in the processing algorithm. The LSWG considered the recommended changes to have little or no effect on the geophysical data and consequently approved them for implementation starting with 1988 data processing. When the processing of the entire historical data (1981 – 1993) is complete, the Processing Team will reprocess the 1987 data to maintain absolute consistency in processing procedures for the entire data set.
The changes made in processing and the data period in which the changes and improvements ARE NOT present are

Modification of the relative azimuth representation	Jan. 1,1987 – Dec. 31, 1987
Error correction in the CLAVR algorithm	Jan. 1,1987 – Dec. 31, 1987
Error correction in the thermal calibration algorithm	Jan. 1,1987 – Oct. 16, 1987
Modification of ozone input data in the atmospheric correction algorithm	Jan. 1,1987 – Dec. 31, 1987
Modification of the compositing algorithm	Oct. 1,1986 – Oct. 31, 1988

A. Relative Azimuth Representation Change

When processing commenced with the 1987 data, the relative azimuth value at nadir (solar azimuth minus sensor azimuth) was set to a value equal to the solar azimuth because the relative azimuth is undefined at nadir (sensor zenith is 0). On average, about 0.016% of the land pixels (plus or minus 3700 pixels/day) are affected in a daily image and less in a 10-day composite. The incorrect nadir relative azimuths have been run through a radiative transfer model and it was determined that the effect on the reflectance value is less than 0.08% in the reflectances and even smaller in the NDVI. The processing software was modified to set the relative azimuth angle at nadir to missing. As a result, the atmospheric correction for these pixels is not calculated and no NDVI is calculated or binned. Second, the relative azimuth is no longer calculated as the absolute value of the solar azimuth angle minus the sensor azimuth angle, but rather is represented by a full 360 degrees. During processing of the 1987 data, if the difference of solar azimuth and relative azimuth is negative (e.g., –20 degrees) the absolute value was taken rather than setting the value to 340 degrees. About one-quarter of the land pixels are affected by this, but the exact locations vary by season. Only those pixels east of nadir and north of the latitude of solar declination are affected. This change affects only the relative azimuth angle data layer in the Daily and Composite Data Set and was corrected beginning with 1988 data processing. Since in the atmospheric correction software the cosine of the realtive azimuth was taken, this change did not affect the reflectances or the NDVI.

B. CLAVR Algorithm Error Correction

A few errors were found in the CLAVR algorithm as implemented in the Pathfinder system. The most significant change is that the Reflectance Gross Cloud Test rarely flagged data as cloudy in the 1987 data set. This was due to a coding error, which has been corrected beginning in 1988 and, as such, interannual comparisons of this layer SHOULD NOT USE

DATA FROM 1987. Furthermore, please be reminded that the CLAVR data layer is an unvalidated, experimental research data layer and the Pathfinder Land Science Working Group recommends that users be cautious about using this layer for data applications and analysis.

C. Thermal Calibration Error Correction

A correction to the "negative radiance from space" was applied incorrectly. These corrections, which change occasionally over the life of the instrument, come from a table provided by NOAA. Only data from January 1, 1987, through October 16, 1987, are affected.

The processing software was corrected, and several days from 1987 were tested to determine the magnitude of the correction. On average, the change in brightness temperature in Channel 4 was 0.71 K and the change in Channel 5 was 0.41 K. Most of the changes occur in cold cloud tops with warm brightness temperatures being essentially unchanged. This change does not affect the NDVI, visible, or geometry layers. Cold areas may have a great enough change in brightness temperature that a CLAVR threshold may flag pixels that would not have been flagged otherwise, but this would be a very rare case.

In addition, it should be noted that during the period of mid-March 1987 and mid-October 1987 there were several days in which there was no thermal calibration information. This is not a processing error but rather a result of communications or level 0 processing errors when the 1b orbital data were generated. The quality control comments indicate those images in which this occurs. The visible channels are fine in these time periods; however, caution should be used when using thermal data from composites or daily images from these time periods.

D. Ozone Data Modification

The ozone data used in calculating the ozone absorption term of the atmospheric correction comes from the Total Ozone Mapping Spectrometer (TOMS) data set. The TOMS data files used in the Pathfinder processing are monthly files containing all daily data. During processing of the 1987 data, data were produced using the ozone from the first day of the each month. Beginning with 1988 data processing, the atmospheric correction for each day uses the ozone data for that day.

Data were processed from several seasons to determine the effect on the data set. On average, Channel 1 changed by 0.02% reflectance and Channel 2 changed by 0.13% and NDVI changed by 0.004. The change to using daily ozone is an improvement in the data set, but the overall effect is small enough that interannual comparisons can still be made.

E. Compositing Algorithm Change

The quality control (QC) flag layer of the Pathfinder data set contains useful and important information about a variety of data conditions. Most of the conditions flagged are simply for information (e.g., the interpolated data flag) and some indicate a condition with minimal impact on the data (e.g., the Channel 1 and 2 nonstandard flag, which indicates that a default ozone value was used).

One exception to this is the QC flag that indicates that a NOAA QC flag was set in the input data (Pathfinder QC flag 16). This can be caused by a number of conditions described in the "NOAA Polar Orbiter Data User's Guide" (Kidwell, 1991). Some of these conditions, such as bad or unavailable calibration data, result in data values that are out of range or are obviously incorrect. However, in some cases NOAA QC flags on the input data are simply for information and do not indicate error conditions in the data.

When bad data (extremely high or extremely low values) are present in the daily data, processing the 8–11 days may include the bad data in the composite. This was changed so that during compositing all data with a Pathfinder QC flag indicating that a NOAA QC flag was set are filtered out. This means that those composites from October 1, 1986, through October 31, 1988, may contain data that were flagged with a NOAA QC flag that may be incorrect. Composites from time periods outside this range (10/1/86 – 10/31/88) will not include any data with a NOAA QC flag; however, there are many cases of extremely high or low data values which result from bad digital counts or calibration in the GAC orbits and for which there is no readily available detection method. These data may be present in composites even when all data with NOAA QC flags have been removed. The Pathfinder processing team plans to regenerate the composites from the 10/1/86 – 10/31/88 time period and update the archive. Users will be informed when this update has occurred. The timing for this archive update is uncertain. Users are advised, in the interim, to carefully check the 10-day composite data and verify that this problem has a minimal effect. Note that this change involves only the Composite Data Set, the daily data are not affected.

Appendix E - NCSA Hierarchical Data Format (HDF) Basics

An important objective of the Earth Observing System Data and Information System (EOSDIS) is to enable scientists to acquire large Earth science data sets from the Distributed Active Archive Centers (DAACs) and many other sources via the online computer network for interdisciplinary studies. Some transportable data formats were recently developed, such as the Hierarchical Data Format (HDF) from the National Center for Supercomputing Applications (NCSA) at the University of Illinois, and the Network Common Data Format from the Unidata Program Center at the University Corporation for Atmospheric Research, Boulder. EOSDIS chose HDF for all its data handling operations.

This appendix presents excerpts from the "Hierarchical Data Format (HDF) User's Guide" by NCSA, University of Illinois at Urbana-Champaign, 1990, with permission of the authors.

What Is Hierarchical Data Format?

HDF is a multiobject file structure designed to facilitate data sharing among people, projects, and machines on a network (see Figure E-1). HDF was created at NCSA to serve the needs of diverse groups of scientists working on supercomputing projects of many kinds.

Why Was HDF Created?

Scientists commonly generate and process data files on several different machines, use various software packages to process files, and share data files with others who use different machines and software. Also, the mixture of information that scientists need to work with often varies from one file to another, even for the same application. Files may be conceptually related but physically separated; e.g., some data may be dispersed among different files, some in program code, and some in the minds of various users.

HDF addresses these problems by providing a general purpose file structure that

- makes it possible for programs to obtain information about the data in a file from the file itself, rather than from another source
- lets you store different mixtures of data and related information in different files, even when the files are processed by the same application program
- standardizes the formats and descriptions of many types of commonly used data sets, such as raster images and scientific data
- encourages the use of a common data format by all machines and programs that produce files containing a specific data set
- can be adapted to accommodate virtually any kind of data by defining new tags or new combinations of tags.

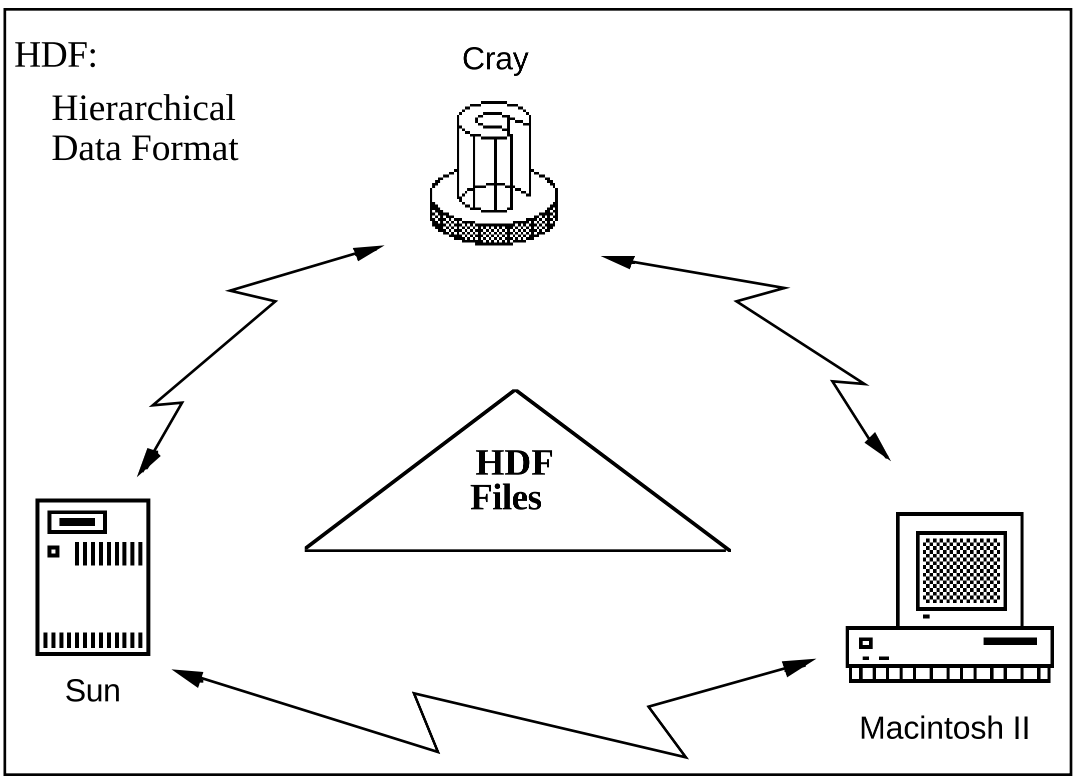


Figure E-1. HDF: A File Format for Scientific Data in a Distributed Environment

HDF files are self-describing. For each data object in an HDF file, there are predefined tags that identify such information as the type of data, the amount of data, their dimensions, and their location in the file.

The self-describing capability of HDF files has important implications for processing scientific data. It makes it possible to fully understand the structure and contents of a file just from the information stored in the file itself. A program written to interpret certain tag types can scan a file containing those tag types and process the corresponding data. Self-description also means that many types of data can be bundled in an HDF file. For example, it is possible to accommodate symbolic, numeric, and graphic data in one HDF file.

Related items of information about a particular type of data are grouped into sets, such as raster image sets and scientific data sets. Each set defines an application area supported by HDF. Additional sets can be defined and added to HDF as needs arise.

Figure E-2 shows a conceptual view of an HDF file containing a scientific data set. The actual two-dimensional array of data is only one element in the set. Other elements include the number of dimensions (rank), the sizes of the dimensions, identifying information about the data and axes, and scales (ranges) for the axes.

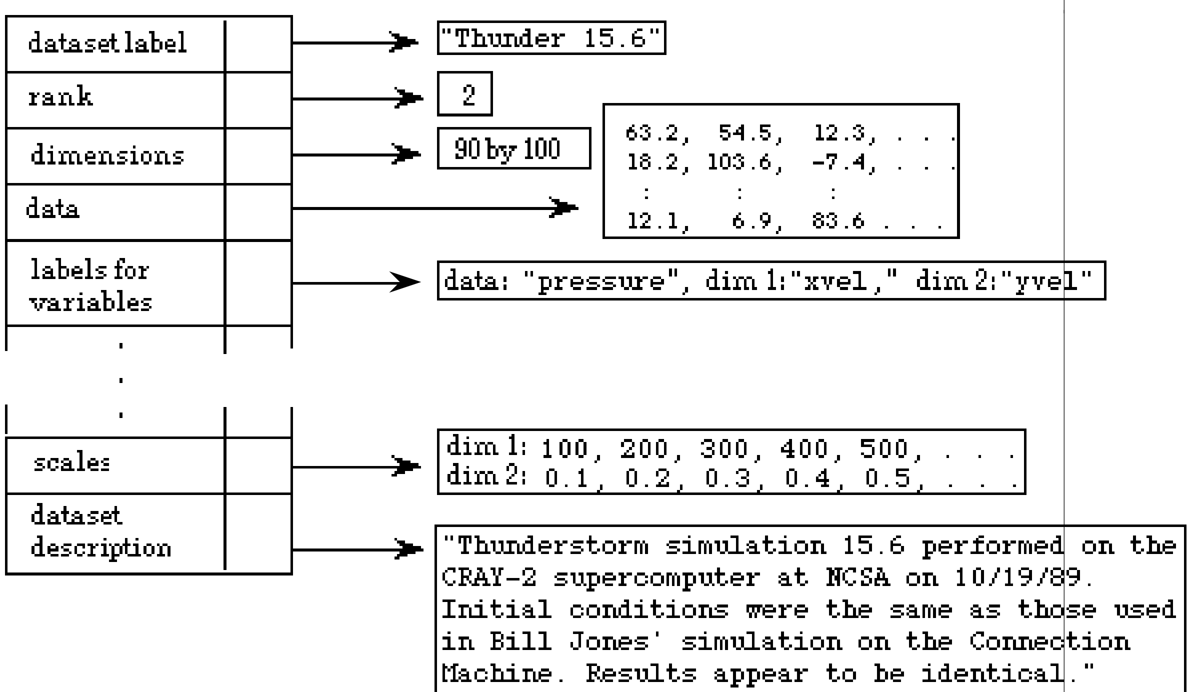


Figure E-2. HDF File with Scientific Data set

NCSA HDF Application Software

NCSA HDF application software currently comes in three forms: (1) NCSA scientific visualization tools that read and write HDF files, (2) calling interfaces that let you read and write HDF files from within a FORTRAN or C program, and (3) command-line utilities that operate directly on HDF files.

The integration of these types of software in the computing environment at NCSA is illustrated in Figure E-3. Visualization tools such as NCSA CompositeTool and NCSA DataScope read and write HDF files. Calling interfaces in the HDF library (libdf.a) let you read and write HDF files from your programs. And utilities such as **r8tohdf** let you operate on HDF files at the command level.

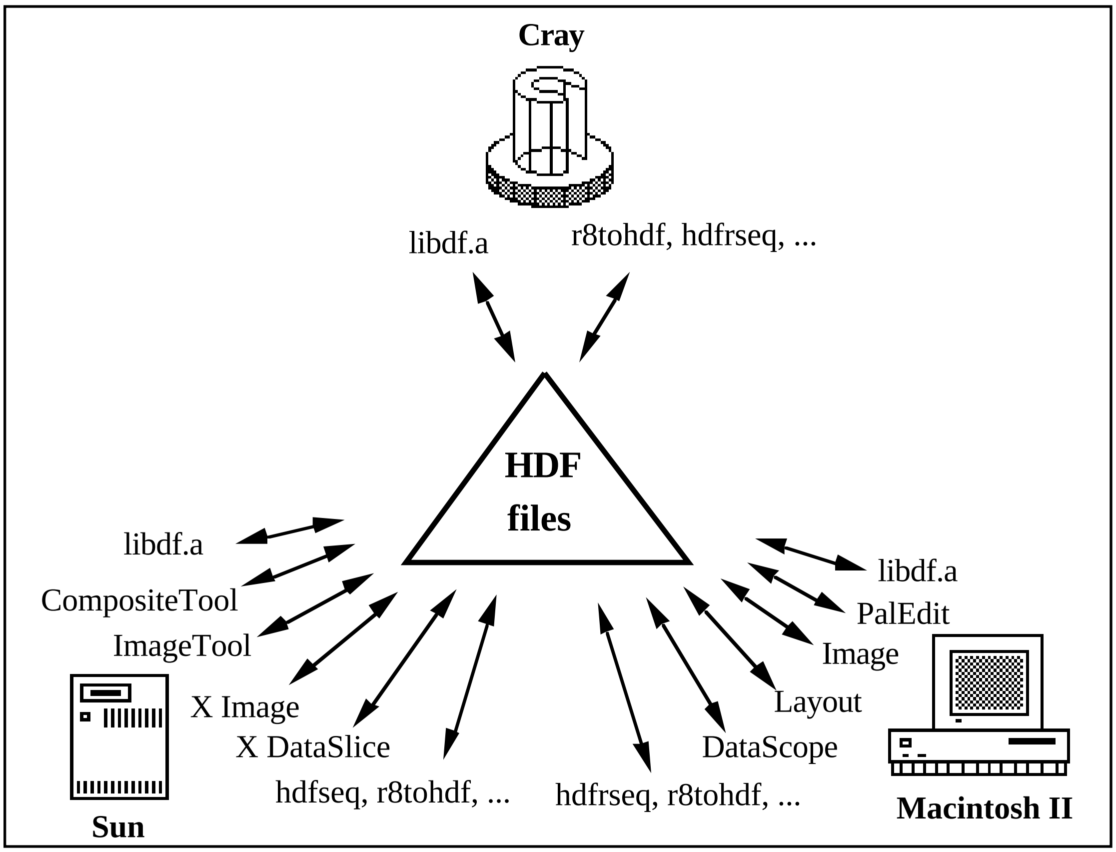


Figure E-3. HDF Software in an Integrated Computing Environment

NCSA Scientific Visualization Software and HDF

The use of HDF files guarantees the interoperability of the scientific visualization tools at NCSA. Some tools operate on raster images, some operate on color palettes, some use images, color palettes, data and annotations, and so forth. HDF provides the range of data types that these tools need, in a format that lets different tools with different data requirements operate on the same files without confusion.

HDF Calling Interfaces

To minimize the amount of knowledge you need to have about HDF, calling interfaces are being developed for specific types of applications, such as storage and display of raster images or scientific data archiving. A calling interface is a library of routines that can be called from an application program for storing and retrieving information, including raw data, from a particular type of HDF file.

Different applications typically require different interfaces. Consequently, NCSA HDF provides FORTRAN and C calling interfaces for storing and retrieving 8- and 24-bit raster images, palettes, scientific data, and annotations. These interfaces are mutually compatible, and user programs can combine calls to routines in different interfaces when they need to store different kinds of data in the same file.

In some rare cases, an application may require the use of a combination of routines from different interfaces. Just as it is possible to define new HDF tags, it is also possible to build new interfaces by combining routines from two or more existing interfaces.

HDF files tend to be used on several different machines, and HDF interfaces developed at NCSA are implemented on as many machines as possible. An important goal in the development of NCSA HDF user interfaces is to eliminate the necessity of changing program code when moving an application from one machine to another.

HDF Utilities

The HDF command line utilities are application programs that can be executed by entering them at the command level, just like other UNIX commands. They make it possible for you to perform, at the command level, common operations on HDF files for which you would normally have to write your own program. For example, the utility **r8tohdf** is a program that takes a raw raster image from a file and stores it in an HDF file in a raster image set.

The HDF utilities provide capabilities for doing things with HDF files that would be difficult to do under your own program control. For example, the utility **hdfseq** takes a raster image from an HDF file and displays it immediately on a Sun-3 console.

Getting Started With HDF

“Public HDF Directories on NCSA Computers,” contains a list of machines at NCSA that have HDF libraries available to all users. If you do not have access to a machine that already has an HDF library, you will need to install it yourself or have your system administrator install it. Although procedures for installing the HDF library vary from one system to another, the basic steps are the same in all cases.

First, you need to get the software from NCSA (see How To Get HDF) or elsewhere. You might get the actual precompiled library, in which case you could load it into your machine into an appropriate directory. If instead you get the source code for the HDF library, you first have to compile the source code into a linkable library.

Some HDF routines are command line utilities, which means that you simply execute them by typing their names in as commands, using the appropriate parameters. Other routines include those that you call from within C or FORTRAN programs. When you write programs that call these routines, you just link the library to your program at compile time.

Detailed information on how to install and use HDF on specific systems can be found in the documentation that comes with the system-specific versions of the software.

Transferring HDF Files

HDF files are binary files, so any transfer protocol that transfers binary files without changing them can be used to transfer HDF files.

Many HDF users use FTP to transfer HDF files. If you use FTP, switch to binary mode when transferring HDF files.

If you use NCSA **telnet** and you wish to transfer an HDF file to or from a Macintosh, you must pay special attention to whether or not to enable the "Macbinary" option. There are two cases to consider.

- If the HDF file is not from a Macintosh application (e.g., it is a normal HDF file generated by your FORTRAN or C program), then be sure to turn Macbinary mode off before performing the transfer.
- If the HDF file corresponds to a Macintosh application (e.g., NCSA Layout, NCSA DataScope, etc.), and you want to transfer it so that it can be accessed from a Macintosh application on another Mac, then be sure to turn Macbinary mode on before performing the transfer.

How To Get HDF

You may obtain NCSA software via FTP, an archive server, or U.S. mail. Instructions for doing so are provided below.

FTP

If you are connected to Internet (NSFNET, ARPANET, MILNET, etc.) you may download NCSA HDF software, documentation, and source code at no charge from an anonymous file transfer protocol (FTP) server at NCSA. If you have any questions regarding this procedure or whether you are connected to Internet, consult your local system administration or network expert.

1. Log onto a host at your site that is connected to the Internet and is running software supporting FTP
2. Invoke FTP on most systems by entering the Internet address of the server:

   ftp ftp.ncsa.uiuc.edu or **ftp 128.174.20.50**
3. Log in by entering **anonymous** for the name.
4. Enter your local login name for the password.
5. Enter **cd HDF** to move to the HDF directory.
6. Enter **get README.FIRST** to transfer the instructions (ASCII) to your local host.
7. Enter **quit** to exit FTP and return to your local host.
8. Review the README.FIRST file for complete instructions concerning the organization of the FTP

directories and the procedure you should follow to download the README files that contain further information on how to get and compile the most recently released version of HDF for your machine and operating system and to determine which files to transfer to your home machine.

Your login session should resemble the sample presented below, where the remote user's local login name is "Smith."

```
harriet_51% ftp ftp.ncsa.uiuc.edu
        Connected to zaphod.
        220 zaphod FTP server (Version 4.173 Tue
        Jan 31 08:29:00 CST 1989) ready.
Name (ftp.ncsa.uiuc.edu: smith): anonymous
        331 Guest login ok, send ident as password.
Password: smith
        230 Guest login ok, access restrictions apply.
ftp> cd HDF
        250 CWD command successful
ftp> get README.FIRST
        200 PORT command successful.
        150 Opening ASCII mode data connection for
        README.FIRST (10283 bytes).
        226 Transfer complete.
        local: README.FIRST remote:
        README.FIRST
        11066 bytes received in .34 seconds (32
        Kbytes/s)
ftp> quit
        221 Goodbye.
harriet_52%
```

NCSA HDF documentation, program, and source code are now in the public domain. You may copy, modify, and distribute these files as you see fit.

Archive Server

To obtain NCSA software via an archive server,

1. E-mail a request to

   archive-server@ncsa.uiuc.edu
2. Include in the subject or message line the word help
3. Press **RETURN**
4. Send another e-mail request to

   archive-server@ncsa.uiuc.edu
5. Include in the subject or message line the word index
6. Press **RETURN**.

For example, if you use the UNIX mailing system, your login session should resemble the following sample, where user entries are indicated in boldface type.

yoyodyne_51% **mail archive-server@ncsa.uiuc.edu**
Subject: **help**
EOT
Null message body; hope that's ok
yoyodyne_52% **mail archive-server@ncsa.uiuc.edu**
Subject: **index**
EOT
Null message body; hope that's ok

The information you receive from both the help and index commands will give you further instructions on obtaining NCSA software. This controlled-access server will e-mail the distribution to you one segment at a time.

Appendix F - Tape Access

Data are available on 4 mm (DAT), high- or low-density 8 mm (Exabyte), and 6250 bpi 9-track tapes. Tapes are created with UNIX utilities **dd** and **tar** on a Silicon Graphics 440 system. The no-swap device and a block size of 63.5 KB are used. Tapes may be requested in **dd** or **tar** file format. By default, the data are archived and distributed in compressed format; however, they are available in uncompressed format by special request. Each tape distributed by the Goddard DAAC contains printed paper labels with the names of the files it contains in the order they were written. Files are compressed using the standard UNIX **compress** command, indicated by a **.Z** appended to the data file name, and should be uncompressed using the UNIX **uncompress** command.

The same utility that was used to create the tape, **dd** or **tar**, must be used to access the data. If compressed files are requested, the data are compressed before the dd or tar utility is executed. Therefore, compressed data should be uncompressed after copying it from tape to local disk.

Data on magnetic tape are written in either standard UNIX **tar** or **dd** format on an SGI computer. The tape block size is 63.5 KB, which translates to a blocking factor of 127.

To read a **tar** format tape on a computer with a UNIX operating system, use the command

tar xvfb <filename> 127

where xvbf are tar command key arguments:

x indicates that the data are to be read from tape
v requests verbose output, i.e., processed file names will be listed
b states that a blocking factor is specified
f states that an archive name is specified
127 is the blocking factor

The fields in < > are system specific and may specify a device, such as a tape drive, or a file directory. The specific parameters depend on your local workstation configuration (e.g., this will be **/dev/8mm1nr** if you read the tape off the 8mm1 tape drive on the DAAC computer with the "no rewind" option).

To read a tar format tape on a VAX computer with a VMS operating system, you will need vmstar public domain software. Use the command

mount /FOREIGN/RECORD=512/BLOCK=65024 <tape drive>

vmstar xvf <tape drive>

To get vmstar, log into the Western Kentucky University anonymous FTP server using the command

ftp ftp.wku.edu

Log into the "fileserv" directory using the command

cd vms/fileserv

and download the file VMSTAR.ZIP to your workstation using the FTP **get** or **mget** commands. Then follow the instructions for installation and unpacking in the README files.

To read a **dd** format tape on a computer with a UNIX operating system, use the command

dd if=<dev> of=<filename> bs=65024

where if=<dev> specifies the tape drive with "no rewind option"

of=<filename> specifies the desired output file name

bs=65024 indicates the block size in bytes

To read **dd** format Pathfinder tapes on a VAX computer with a VMS operating system, simply use the **copy** command. A public domain software, lzw, is available to read the UNIX compressed data format on VAX VMS. To read compressed data files, you must first use **vmstar** or the **copy** command to unload the data from tape. Then issue the **lzw** command.

REFERENCES

Baldwin, D., and W. J. Emery. 1993. Systematized approach to AVHRR navigation. Annals of Glaciology, 17:414–420.

Bowman, K.P., and A.J. Krueger. 1985. A global climatology of total ozone from the Nimbus–7 Total Ozone Mapping Spectrophotometer. Journal of Geophysical Research, 90:7967–7976.

Brown, S.A., M. Folk, G. Goucher, and R. Rew. 1993. Software for portable scientific data management. Computers in Physics, 7:304–308.

Brown, O.W., J.W. Brown, and R.H. Evans. 1985. Calibration of Advanced Very High Resolution Radiometer observations. Journal of Geophysical Research, 90:11667–11677.

Che, N., and J.C. Price. 1992. Survey of radiometric calibration results and methods for visible and near-infrared channels of NOAA–7, –9 and –11 AVHRRs. Remote Sensing of the Environment, 41:19–27.

Gordon, H.R., J.W. Brown, and R.H. Evans. 1988. Exact Rayleigh scattering calculations for use with the Nimbus-7 coastal zone color scanner. Applied Optics, 27:2111–2122.

Goward, S.N., B. Markham, D.G. Dye, W. Dulaney, and J. Yang. 1991. Normalized difference vegetation index measurements from the Advanced Very High Resolution Radiometer. Remote Sensing of the Environment, 35:257–277.

Goward, S.N., D.G. Dye, S. Turner, and J. Yang. 1993. Objective assessment of the NOAA global vegetation index data product. International Journal of Remote Sensing, 14:3365–3394.

Holben, B.N. 1986. Characteristics of maximum-value composite images from temporal AVHRR data. International Journal of Remote Sensing, 7:1417–1434.

Jackson, R.D., P.N. Slater, and P.J. Pinter. 1983. Discrimination of growth and water stress in wheat by various vegetation indices through clear and turbid atmospheres. Remote Sensing of the Environment, 15:187–208.

James, M.E., and S.N.V. Kalluri. 1994. The Pathfinder AVHRR land data set: An improved coarse resolution data set for terrestrial monitoring. International Journal of Remote Sensing, (in press).

Justice, C.O., B.L. Markham, J.R.G. Townshend, and R.L. Kennard. 1989. Spatial degradation of satellite data. International Journal of Remote Sensing, 10:1539–1561.

Kaufman, Y.J., and B.N. Holben. 1993. Calibration of the AVHRR visible and near-IR bands by atmospheric scattering, ocean glint, and desert reflection. International Journal of Remote Sensing, 1:21–52.

Kidwell, K.B. 1990. Global Vegetation Index User's Guide. NOAA/NESDIS, National Climatic Data Center, Washington D.C.

Kidwell, K. 1991. NOAA Polar Orbiter Data User's Guide. NCDC/SDSD, National Climatic Data Center, Washington, DC.

NCSA. 1990. Hierarchical Data Format (HDF) Users Guide. National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana-Champaign.

NGDC. 1993. ETOPO5 5-minute gridded elevations/bathymetry for the world. Data announcement 93-MGG-01, National Geophysical Data Center, Boulder.

Patt, F., and W. Gregg, 1994. Exact closed-form geolocation algorithm for Earth survey sensors. International Journal of Remote Sensing, (in press).

Rao, C.R.N. 1993a. Non-linearity corrections for the thermal infrared channels of the Advanced Very High Resolution Radiometer: assessment and recommendations. NOAA Technical Report NESDIS-69, NOAA/NESDIS, Washington, DC.

Rao, C.R.N. 1993b. Degradation of the visible and near-infrared channels of the Advanced Very High Resolution Radiometer on the NOAA–9 spacecraft: Assessment and recommendations for corrections. NOAA Technical Report NESDIS-70, NOAA/NESDIS, Washington, DC.

Sellers, P.J., Y. Mintz, Y.C. Sud and A. Dalcher, 1986. A simple biosphere model (SiB) for use with general circulation models. Journal of Atmospheric Sciences, 43:505–531.

Smith, G. R., R.H. Levin, P. Abel and H. Jacobowitz. 1988. Calibration of the solar channels of NOAA–9 AVHRR using high altitude aircraft measurements. Journal of Atmos. Ocean. Technology, 5:631–639.

Staylor, W.F. 1990. Degradation rates of the AVHRR visible channel from the NOAA–6, 7 and 9 spacecraft. Journal of Atmospheric and Oceanic Technology, 7:411–423.

Steinwand, D.R., J.A. Hutchinson, and J.P. Snyder, 1992. Map projections from global and continental data sets, and an analysis of distortion caused by reprojection. USGS/EDC contract report, EROS Data Center, Sioux Falls.

Steinwand, D.R. 1994. Mapping raster imagery to the Interrupted Goode Homolosine Projection. International Journal of Remote Sensing, (in press).

Steyn-Ross, D.A., M. L. Steyn-Ross, and S. Clift. 1992. Radiance calibrations for Advanced Very High Resolution Radiometer infrared channels. Journal of Geophysical Research, 97:5551–5568.

Stowe, L.L., E.P. McClain, R. Carey, P. Pellegrino, G. G. Gutman, P. Davis, C. Long, and S. Hart. 1991. Global distribution of cloud cover derived from NOAA/AVHRR operational satellite data. Advances in Space Research, 3:51–54.

Strebel, D.E., D.R. Landis, J.A. Newcomer, D. van Elburg-Obler, B.W. Meeson, and P.A. Agbu. 1992. Collected data of the First ISLSCP Field Experiment, Volume 2: Satellite Imagery 1987–1989. Published on CD-ROM by NASA, Goddard Space Flight Center, Greenbelt.

Tanre, D., B.N. Holben, and Y. Kaufman. 1992. Atmospheric Correction Algorithm for NOAA-AVHRR Products: Theory and Application. IEEE Transactions on Geoscience and Remote Sensing, 30:231–248.

Townshend, J.R.G. 1992. Improved global data for land applications: a proposal for a new high resolution data set. International Geosphere Biosphere Program report no. 20. Stockholm.

Tucker, C.J. 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of the Environment, 8:127–150.

Tucker, C. J., W. W. Newcomb, S. O. Los, and S. D. Prince. 1991. Mean and inter-year variation of growing-season normalized difference vegetation index for the Sahel 1981–1989. International Journal of Remote Sensing, 12:1113–1115.

Vermote, E., D. Tanre, and M. Herman. 1990. Atmospheric effects on satellite imagery, correction algorithms for ocean color or vegetation monitoring. International Society for Photogrammetry and Remote Sensing, 28:46–55

Weinreb, M.P., G. Hamilton, S. Brown, and R.J. Koczor. 1990. Nonlinearity corrections in calibration of Advanced Very High Resolution Radiometer infrared channels, Journal of Geophysical Research, 95:7381–7388.

ACRONYM List

ANSI	American National Standards Institute
ASCII	American Standard Coded Information Interchange
AVHRR	Advanced Very High Resolution Radiometer
BISE	Best Index Slope Extraction
bpi	bits per inch
CDF	Common Data Format
CD-ROM	Compact Disk–Read Only Memory
CLAVR	Clouds for AVHRR
CWG	Calibration Working Group
DAAC	Distributed Active Archive Center
DAT	Digital Audio Tape
DMSP	Defense Meteorological Satellite Program
EDC	EROS Data Center
EOS	Earth Observing System
EOSDIS	EOS Data and Information System
EROS	Earth Resources Observations System
ETOPO5	Earth Topographic 5-Minute Grid
FTP	File Transfer Protocol
GAC	Global Area Coverage
GCDC	Global Change Data Center
GCM	Global Climate Model
GCTP	General Cartographic Transformation Package
GIMMS	Global Inventory Monitoring and Modeling System
GOES	Geostationary Operational Environmental Satellite
GVI	Global Vegetation Index
HDF	Hierarchical Data Format
HP	Hewlett Packard
HRPT	High Resolution Picture Transmission
IDL	Interactive Data Language
IEEE	Institute of Electrical and Electronics Engineers
IGBP	International Geosphere Biosphere Program
IMS	Information Management System
ISCCP	International Satellite Cloud Climatology Program
ISLSCP	International Satellite Land-Surface Climatology Project
JRC	Joint Research Centre

KB	Kilobyte
LAC	Local Area Coverage
LSWG	Land Science Working Group
MB	Megabyte
MTV	Monitoring of Tropical Vegetation
NASA	National Aeronautics and Space Administration
NCDC	National Climatic Data Center
NCSA	National Center for Supercomputing Applications
NDVI	Normalized Difference Vegetation Index
NESDIS	National Environmental Satellite Data and Information Service
NGDC	National Geophysical Data Center
NOAA	National Oceanic and Atmospheric Administration
NODC	National Oceanographic Data Center
NORAD	North American Air Defense Command
PAL	Pathfinder AVHRR Land
PODUG	Polar Orbiter Data User's Guide
QC	Quality Control
RIS	Raster Image Set
SDS	Scientific Data Set
SDSD	Satellite Data Services Division
SeaWiFS	Sea-viewing Wide Field-of-view Sensor
SGI	Silicon Graphics Inc.
SMMR	Scanning Multichannel Microwave Radiometer
SSM/I	Special Sensor Microwave/Imager
TBM	Terabit Memory
TCP/IP	Transmission Control Protocol/Internet Protocol
TIROS	Television and Infrared Operational Satellite
TOMS	Total Ozone Mapping Spectrometer
TOVS	TIROS Operational Vertical Sounder
USGS	United States Geological Survey
USO	User Services Office
WORM	Write Once Read Many